

FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTTCAGATCTGCTCGGTAGA
 CCTGGTGCACCACCACCATGTTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG
 GTTTTCCACCCAGCTTTCACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA
 ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA
 GAACTGGCCAAGAACTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT
 GATCAGATGGGAAGATGGTTTGTGCTGGAGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA
 CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT
 ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAAACAGCT
 TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACTTCATGATGAGAGGCTCTTG
 GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC
 CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT
 GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCTCTTCTCATCAGAGCTGCATGGTACAC
 AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCAGTGAAAAGTTTCTGA
 ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCCTCATTGGGATCTATG
 TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT
 AGTTCCTTTTTCAGCATGTTCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT
 CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT
 ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAAATG
 AAGTGA CT CAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTAATGGGGCAGATATGC
 ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA
 TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTCTGGAGAATAAATGCAGT
 AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAAATGTT
 TTGGTGAATGTGAAAATAAAGTTTGTGTGATGAGAATGTAAGTCTTTTTTCTACTTTAAAA
 TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT
 GCAGAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTTGGTAAAGGGACCAGAGAGAAG
 GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA
 GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAACAAGTGGTCATTGTTACATTCATTT
 GCTGAACTTAACAAAACCTGTTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG
 CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC
 ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA
 ATAAC TTTAAA ACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG
 AATACAAACAGTATACTCATG

030604 1406550

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWVAVAGGAAVGLGALCYGLGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTLGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFMSMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

099044-44504
T05T "T403660

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCCGGTGGCGCCACGTCCGCCCGTCTCCGCCTTCTGCAT
CGCGGCTTCGGCGGCTTCCACCTAGACACCTAACAGTCGCGGAGCCGCGCGCTCGTGAGGG
GGTCGGCACGGGGAGTCGGGCGGTCTTGTCATCTTGGCTACCTGTGGGTCTGAAG**ATG**TCGG
ACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCGCGGCCACCGTC
GCCGTGCCCTTGGTCGGCAAACCTCGGCCTCATCAGCCCGGCCCTACCTCTTCCTCTGGCCCGA
AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG
GTCCAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA
CTTGAAACAGGAGCTTTTGATGGGAGGCCAGCAGACTATTTATTCATGCTCCTCTTTAACTG
GATTTGCATCGTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCTCTGATCA
TGTCAGTACTTTATGTCTGGGCCCAGCTGAACAGAGACATGATTGTATCATTTTGGTTTGGA
ACACGATTTAAGGCCTGCTATTTACCCTGGGTATCCTTGGATTCAACTATATCATCGGAGG
CTCGGTAATCAATGAGCTTATGGAAATCTGGTTGGACATCTTTATTTTTCTTAATGTTCA
GATACCCAATGGACTTGGGAGGAAGAAATTTCTATCCACACCTCAGTTTTTGTACCGCTGG
CTGCCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC
TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTCGACTTGGAGACC
AG**TGA**AGGGGCGGCCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCTCCAGTGCTGGGTG
CACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC
AGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCCACAAAGTTTCACGAT
TCTCATTCAAGTCCTTACTGTGTGAAGAACAAATACCAACTGTGCAAATTGCAAACTGAC
TACATTTTTTGGTGTCTTCTCTTCTCCCCTTTCCGTCTGAATAATGGGTTTTAGCGGGTCTCT
AATCTGCTGGCATTGAGCTGGGGCTGGGTACCAAACCCCTTCCCAAAGGACCTTATCTCTT
TCTTGACACATGCCTCTCTCCCACCTTTCCCAACCCCCACATTTGCAACTAGAAAAAGTTG
CCCATAAAATTGCTCTGCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGTC
ACAACAATCATATTCACGTATTTTCCCCTTTTGGTGGCAGAAGTGTACCAATAGGGGGGAG
AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATTGCTTCGACTGACATCCGTTGTT
AACCGTTTGCCACTCTTCAGATATTTTTTATAAAAAAAGTACCACTGAGTTCATGAGGGCCA
CAGATTGGTTATTAATGAGATACGAGGGTGGTGTGGGTGTTTGTTCCTGAGCTAAGTGA
TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTAGGTTTAAACCATGGGGGATGCACCCC
TTTGCGTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTGCTTTGTGT
TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG
ATTCCCATTTCAATCTCATTCTGGATATGTGTTCAATTGAGTAAAGGAGGAGAGACCCCTCATA
CGCTATTTAAATGTCACCTTTTTTGCCTATCCCCCGTTTTTTGGTCATGTTTCAATTAATTGT
GAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTTAAAGCTAATGTAAGCACATCTA
AGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA
TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTTT
TCGTAGGTGGGCTTTTCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCA
TGGCTTTTTCACACAGTTATTTTATTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT
ATTGAGTGGCTGTACACTTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTTGATCAGTTTCTT
TTCAGGAAACATTTGTGCTCTAACAGTATGACTATTCTTTCCCCCACTCTTAAACAGTGTGAT
GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG
TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA
TGTTTTGTGTTTCATCTGTGGCCACAATAAAGTTTACTTGTAAAATTTTAGAGGCCATTACT
CCAATTATGTTGCACGTACACTCATTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATA
TTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGTACCCTCTTACCAGTCAGCTGCCTGCGAG
CAGTCATTTTTTCTTAAAGGTTTACAAGTATTTAGAACTTTTCAGTTCAGGGCAAAATGTTT
ATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT
ATGTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTGTGACTTGACAAGGCAAAACA
TGACAGTGGATTCTCTTACAAATGGAAAAAATAATCCTTATTTTGTATAAAGGACTTCCC
TTTTTGTAACATAATCCTTTTTTATTGGTAAAAATTGTAAATTAAATGTGCACTTG

0998041.1.1501
T0511.1408660

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF
PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLEMLLFNWICIVITGLAMDMQLLMIP
LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIIGGSVINELIGNLVGHLYFFL
MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL
GDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCCATTAACTGGGTG
GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT
CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC
CCTTTGGGGCGGG**ATG**GCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG
CGGGGTTCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT
GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT
GGCCTGTGTTCCCCTTGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC
ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT
AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC
CATTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA
AAAACATTGAAATGCAGCTGCAAGCCATTTCGAATAATTCAAGAGAGAAATGGTGTATTACCT
GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT
GAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAGGAAAA
AACAGTTATCAGAGGCTAAAACAGAAGAGCCACAGTGCATTCCAGTGAAGCTGCAATAATG
AATAATTCCTAAGGGGATGGTGAACATTTTGCACACCCACCCTCAGAAGTTAAAATGCATTT
TGCTAATCAGTCAATAGAACCCTTTGGGAAGAAAAGTGGAAAGGTCTGAACTTCCTCCCTCC
CACAAAAGGCCTGAAGATTCTTGCTTAGAGCATGCGAGCATTTGAAGGACCAATAGCAAAC
TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA
TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG
GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAG
CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAAACTCAAAGAAGAAGTTATTAATAAG**TA**
ATAATTAAGAACAATTTAACAAAATGGAAGTTCAAATTGTCTTAAAAATAAATTATTTAGTC
CTTACACTG

FIGURE 6

MAAEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP
LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEACTSPLAKTHTSQAILQP
VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGLPDCLTDGSDVVSDLEHEEMKILREVL
RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAIMNNSQGDGEHFAHPPSEVKMHFANQS
IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS
MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEKQTLLKRLLAEKLKEEVINK

N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

N-myristoylation site.

amino acids 273-279

Amidation site.

amino acids 252-256

Cytosolic fatty-acid binding proteins.

amino acids 78-108

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTCTGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAA

FIGURE 8

GCGTGGTTTTTTGTTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTG
 TAGCTTCTCCACGTATGGACCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG
 TCTCAGCTCTAGGATGTGCGTTCTTCCACTAGAAAGCTCTTCTGAGGGAGGTAATTA AAAAAC
 AGTGGA**ATG**GAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCACAATGTATAC
 ATTCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT
 CTGCCAATGAAGAAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTG
 GTGAAGCTAGTTTTCTGTGTGCTTGTGTCAATTCTGTGTTATAAAGAAAGATCATCAAAGTAG
 AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCCTGCCT
 TTCTTTATTTCTGGATAACTTGATTGTCTTCTATGTCCTGTCTATCTTCAACCAGCCATG
 GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTACAGGATAGTGCTGAA
 GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCT
 TGACTGCCGGGACTAAAACCTTTACAGCACAACCTGGCAGGACGTGGATTTTCATCACGATGCC
 TTTTTCAGCCCTTCCAATTCCTGCCTTCTTTTCAGAAGTGAGTGTCCAGAAAAGACAATTG
 TACAGCAAAGGAATGGACTTTTCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTTCAGTC
 ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT
 AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA
 GAACAGCAAACCTCTATTTCTTTGGCATTCTGTTTAAATGGGCTGACTCTGGGCCTTCAGAGGA
 GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTTATGGCCACAGTGCATTTTCAGTAGCC
 CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA
 CATGTTCCATGTCTTGATGGCCCAGGTTACCACTGTCATTATCACAAACAGTGTCTGTCTCTGG
 TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGGAAGCCCCATCAGTCCTTCTCTCTATA
 TTTATTTATAATGCCAGCAAGCCTCAAGTTCGGAATACGCACCTAGGCAAGAAAGGATCCG
 AGATCTAAGTGGAATCTTTGGGAGCGTTCAGTGGGGATGGAGAAGAACTAGAAAGACTTA
 CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTCT**TAA**CTGGTACCCACATAGTTTGCA
 GCTCTCTTGAACCTTATTTTTCACATTTTTCAGTGTGTTGTAATATTTATCTTTTCACTTTGATA
 AACCAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT
 CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA
 ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT
 TTTCTTGGCCTTCAAGCTTCCAAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATAT
 ACACATAGAGATCAATTTGCCAAATATTACAAATCATGTAGTTCTAGTTTACATGCCAAAGT
 CTTCCCTTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT
 CATTTTGAAGTAAAGAGCAACGGGACCCTTTCTAAAAACGTTGGTTGAAGGACCTAAATAC
 CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGCTGAAGAAGCAGT
 TTCTCAGACACAACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT
 AATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA
 AGTCATTTAAACAAGCCACGGTGGGGCTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT
 GTCATTACTCCTGAATTATTACATTTTGGAGAATAAGAGGGCATTTTATTTTATTAGTTACT
 AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC
 CAGATTGTCAGTGAAGCTGATGCCTAGGAACCTTTTAAAGGGATCCTTTCAAAGGATCACTT
 AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC
 AGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAGTGCATGGTATTTTTTCATG
 GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA
 AAATTAGCAAACAAAAGTGAAGTGTCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG
 CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT
 CATTCTCAGAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA
 AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG
 GTGGAAATTTGTAATTAAATAATTATTAAACCT

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLPAMAV
IFS NFSIITALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTGLQRSNRDQIKNCGFFYGHSAFSVALI
FVTAFQGLSVAFILKF LDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESDETF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293,
298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

FIGURE 10

CGTGCCTGCGCAATGGGTGTCGGGTCCGCTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTGTAGCTTCTCCAC
GTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAACAGTGGAATGGAA
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATACATTCTGCTAGG
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG
AAAACAAGTATGATTATCTTCCAACACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT
TTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA
TGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCTGCCTTTCTTTATTTCC
TGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATGGCTGTTATCTTC
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTTCAGGATAGTGCTGAAGAGGCGTCTAAA
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCTTGACTGCCGGGA
CTAAAACTTTA

CGGACGCGTGGGCGGACGCGTGGGCGGAGCGCTGGGGCCGGCTTGCGCTAGCGCGCGGCGGCC
GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT
CCCGTCTACCGAGGTCCCAAGCGGCGTGGCCCGCGGGTTCATGGCCAAAGGAGAAGGCGCCGAG
AGCGGCTCCGCGGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA
GGTGAAGAAAGAACCGAAAAAGAAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG
CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCCTTCAGATCTAC
CTATTGG**ATG**TGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC
CTGGGATGCCATCACAGACCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC
TGGGTGCGCTTATGCCCTGGATCATCTTCTCCACGCCCTGGCCGTCAATTGCCACTTCTCT
ATCTGGTTCTGCCCCGACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCCT
CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA
ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC
AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG
ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATACACATGGCACCACTTCACAC
AGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCAATTGTCTGTATCTATATAATCTG
TGCTGTCACTCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTG
AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTT
ATTACTGGCTTCTCTTACCTCCTTGGCTTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTT
TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT
CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT
GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA
CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC
TACCCTGGTCCATGCTGCCTGATGTCAATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT
GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTTCACCAAGTTTGCCTCTGGAGTGT
ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC
CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG
CTGGGCCTGCTGCTCTTCAAAATGTACCCCATTGATGAGGAGAGGCGGCGGCAGAATAAGAA
GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG
AGCTGGCTAGCATCCTC**TAG**GGCCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG
AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA
CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG
TGGCCTCCTGCCTCCCCTCTGCCTGCCTGTGGGGCCAAAGCCCTGGGGCTGCCACTGTGAATA
TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTTTACAGAGCC
TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTGTATGTATATGTCTGTGAGCTA
TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

FIGURE 12

MWLRWALSLPPSSCLWAEPPGMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG
SCPTSHTARPIGTCSFIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVI
GTAIQGQIVGQADTPCFQDFNSSTVASQSANHHTGTTSHRETQKAYLLAAGVIVCIYIICAV
ILILGVREQREPYEAAQQSEPIAYFRGLRLVMSHGPYIKLITGFLFTSLAFMLVEGNFVLFCT
YTLGFRNEFQNLLLAIMLSATLTIPIWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI
ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG
ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKKAL
QALRDEASSSGCSETDSTELASIL

FIGURE 13

GGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACCTTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTTAACCATTCCTCATCTGGCAGTGGTTCCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCCAT

FIGURE 14

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCTGGTAAGGATTTACAAAAGGTGCAGGT
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGT
GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG
CAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA
GAGAACGTTATCATCAAATTAACAAGGCTGGCCTTGTAAGTTCGTAATACTGAGTTGTTTAGG
ACTTTCTATTGTGGCAAACCTTCAGAAAACAACCCCTTTTGTCTGCACATGTAAGTGGAGCTG
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTTCAGACCATCCTTTCCTACCAAATG
CAGCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG
AGTAAGTGCACCTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT
ACTGCAGCAGAATGGTCTATGTCAATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGA
TTTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCAGAGATATTTGATGAAAGGAT
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAGTTGCTTA
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA
TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFIIFS¹YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLT²YIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRLLSRDI

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 16

CGGACGCTTGGGCNGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTTGAAGCAGT
TACCAAGAATCTTCAACCCTTTCCCACAAAAGCTAATTGAGTACACGTCCTGTTGAGTACA
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA
GTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

FIGURE 17

CCCACGCGTCCGCCCCGCCGTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGG
 CCGGGGTGCGGAGCCGAC**ATG**CGCCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC
 CTTCGCCTTGTACTTGCTGTGACGCGACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGG
 AGGCTGGAGGCAGGTCGCTGTGGTTCCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG
 GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCTGCTCTTCTGCGGCGCCTA
 CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT
 TGTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTTGACCTCGGTGGGTGCCACATGC
 TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCCTGATAAAGT
 GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTTCTTATGTTTTT
 TGAGACTTTTCCCCATGACACCAAACCTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT
 CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTTCATCTGTGT
 GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG
 TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT
 AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA
 CACAT**TGA**TCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA
 TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG
 CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT
 TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT
 GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC
 ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAAT
 ACAAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC
 AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT
 CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPGRRLGSTEEAGGRSLWFPSDLAELRESEVLREYR
 KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSVGATCCYLLSS
 IFGKQLVVSYPDPKVALLQRKVEENRNSLFFFLFLRLFPMTPNWFLNLSAPILNIPVQFF
 FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKSQKHLQ
 LNETSTANHIHSRKDT

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

FIGURE 19

CCGAGGCGGGAGGAGCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTT
CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTTCCGAAC TGCCAGCTCAGAA
TAGGAAAATAA CTTGGGATTTTATATTGGAAGAC **ATG** GATCTTGCTGCCAACGAGATCAGCA
TTTATGACAAACTTT CAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATG
TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC
CCCCCGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT
TGCTCACTGCCTACTTTGTGATTCAACCTTT CAGCCCATTAGCACCTGAGCCAGTGCTTTCT
GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA
GAAGTACATGT CAGAAAATAAGGGAGTTCTCTGCATGGGGGTGATGAAGACAGACCCTTTC
CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTTGCC
AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA
ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC
AGCATTTTTTTGTGCCAGTACCTTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTTCGCCAAG
TGGTGGCGCTGCTTTCTTGAGCGGTGGTTCCCATTTCTTATCCATGGAGGAGACCTCTGAA
CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTCACTCACCTGCCATTTCCAAAAGATG
CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG
ATGCCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA
GTGCCGAAGACATTGT CAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTG
ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT
GGAACCGCTTTCTCAGAACTG **TAG** GAAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA
AAACCAGGTTGAAAGGGGAAAAATAAAAAACAAAACGATGAAACTGCAAAAA

MDLAANEISIIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPPQYPLLIVVY
KVLATLGLILLTAYFVIQPFSPLAPEPVLGSAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL
HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMILLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFLCQYPEATEGFSEGGFFAKWWRCFPERWFPPYPWRRPLNRSQMLRELFPV
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRRCQSVAMP
IEPGDIGYVDTHHWKVYVIARGVQPLVICDGTAFSEL

FIGURE 21

CCACGGTGTCCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTTCAGTTGTGATC
AAGGGACACGTGGTTTCCGAACCTGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAAACTTTCAGAGACTGTTG
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAGAGAAGGCAATTGAAAAATTTATC
AGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCCGCAGTATCCTCTCCTTATAGT
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

CCACACGCGTCCGCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG
CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTG
CCACTGGTGGCGCAGCTGCTAGACCGTGCCATATGAGCCGCTGGGGCTGCAGTGGGGACTGCC
CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA
CACCCGAATGGCGCCACTTCATCGACAAACAGGTACAGCCAAC**CATG**TCCCGAGTTCGAAATG
GACACGTATGCTAAGAGCCACGACCTTATGTCAGGTTTCTGGAATGCCTGCTATGACATGCT
TATGAGCAGTGGGCAGCGGCGCCAGTGGGAGCGCGCCAGAGTCGTCGGGCCTTCCAGGAGC
TGGTGCTGGAACCTGCGCAGAGGCGGGCGCGCTGGAGGGGCTACGCTACACGGCAGTGTCT
AAGCAGCAGGCAACGCGACACTCCATGGCCCTGCTGCATGGGGGGCGCTGTGGCGCCAGCT
CGCCAGCCCATGTGGGGCTTGGGCGCTGAGGGACACTCCCATCCCCCGCTGGAACATGTCCA
GCGCCGAGACATATTACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCCTCAC
CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCTTGACACCCACCGAGGAGGC
CTCACTGCCTCTGGCAGTGACCAAAGAGGCCAAAGTGAGCACCCACCCGAGTTGCTGCAGG
AGGACACAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCGACGAACTG
GATGAGCAGCGTGAGAAGCTGGTGCTGTGCGGCGAGTGCCAGCTGGTGACGGTAGTGCCCTG
GGTCCCAGGGCTGCTGGAGGTCACCACACAGAATGTATACTTCTACGATGGCAGCACTGAGC
GCGTGGAACCCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCCAGCTGCGTGAG
GTCCACCTGCGGCGTTTCAACCTGCGCGCTTCAGCACTTGAGCTCTTCTTTATCGATCAGGC
CAACTACTTCTCACTTCCCATGCAAGGTGGGACGACCCCACTCATCTCCTTAGCCAGA
CTCCGAGACCCCAAGCTGGCCCCATCCCCCCCCATACCCAGGTACGGAACCAAGTGACTCG
TGGCTCCTGCGCCTACGGGCCCCCTCTCAAGGCTACCTAAGCAGCGCTCCCCCAGGAGAT
GCTGCGTGCCCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACCTTCGAGTACT
TGATGCAACTCAACACCATTGCGGGGCGGACCTACAATGACCTGTCTCAGTACCCCTGTGTTT
CCCTGGGTCTCTGAGGACTACGTGTCCCCAACCTTGGACCTCAGCAACCCAGCCGCTCTTCCG
GGACCTGTCTAAGCCCATCGGTGTGGTGAAACCCCAAGCATGCCAGCTCGTGAGGAGAAAGT
ATGAAAGCTTTGAGGACCCAGCAGTGAGGACCAATTGACAAGTTCCTACTATGGCACCCACTCTC
AATGCAGCAGGCGTGATGCACTACCTCATCCGCGTGGAGCCCTTCACCTCCCTGCACGTCCA
GCTGCAAAGTGCCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGGCAGCCTGGC
AGGCACGCCTGGAGAGCCCTGCCGATGTGAAGGAGCTCATCCCGGAATTCTTCTACTTTCTCT
GACTTCTGGAGAACAGAACGTTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGT
AGGCGATGTGGTGCTACCCCCCTGGGGCAGCTCTCTGAGGACTTCATCCAGCAGCACCGCC
AGGCTCTGGAGTCGGAGTATGTGTCTGCACACCTACACGAGTGGATCGACCTCATCTTTGGC
TACAAGCAGCGGGGGCCAGCCGCCGAGGAGGCCCTCAATGTCTTCTATTACTGCACCTATGA
GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTA
TCAGCAACTTTGGGCAGACTCCCTGTACAGTGTCTGAAGGAGCCACATCCACTCGGCTCTCA
GCTGAGGAAGCAGCCCATCGCCTTGACAGCCTGGACACTAACTCACTTAGCATCTTCCAGCA
CCTGGACGAACTCAAGGCATTCTTTCGACAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA
CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC
ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGT
GAGTGGACAAAGCACTGGCAGTGGCCCGGATGGAAAGCTGTCTATTACGCGTGGCCACTGGG
ATGGCAGCCTGCGGGTGACTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCAC
CTTGATGTAGTAACCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCG
GGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTAGGCCTGGCAC
CAAAGCCTGTGCAGGTCCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT
GAACCTGACATGGCTGTGCTGTGATCTGAGGATGGAACATGTGATCATACACACTGTACGCCG
CGGACAGTTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCTTGGACCTATTTTCCACC
TGGCATTTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTTGGGGCC
CAGGTACCTACTCCTTGACCTGTATTAGTCAATGGGAAGTTGCGGGCTTCACTGCCCTT
GGCAGCGCAGCCTACAGCCCTGACGGTGACAGAGGACTTGTGTTGCTGGGCACCGCCAGT
GCGCCCTGCACATCTCCAACATAACACACTGCTCCCGGCGCGCTCCTTGGCCCATGAGG
GTGGCCATCCGCAGCGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCTGGAGGA
TGGCAAGCTCATCGTGGTGGTTCGCGGGGCAGCCCTCTGAGGTGCGCAGCAGCCAGTTTCGCGC
GGAAGCTGTGGCGGTCTCTCGCGGCGCATCTCCAGGTGTCCTCGGAGAGACGGAATACAAC
CCTACTGAGGCGCGCT**TGA**ACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCCGGCAGGCCTG
GCCCGGAGGGCCCCGCCAGAAGTCGGCGGGGAACACCCCGGGGTGGGCAGCCAGGGGGTGA
GCGGGGCCACCCCTGCCAGCTCAGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCG
GGCGGAAGTCCCGCCCTCGCCGGCTGAGGGGGCCGCCCTGAGGGCCAGCACTGGCGTCT

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL
 RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN
 HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAEELETP
 MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP
 LAQLREVHLRFRNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV
 RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL
 SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH
 YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKELEIP
 EFFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPWASSPEDFIQQHRQALESEYVSAHLHEW
 IDLIFGYKQRGPAEEALNVFYYCTYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP
 HPTRLSAEEAAHRLARLDTNPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF
 SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGDSLRTALPRGKLL
 SQLSCHLDVVTCLALDTCGIYLISGSRDTCMVWRLLHQGGLSVGLAPKPVQVLYGHGAAVS
 CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA
 WERPGAQVTYSLHLYSVNGKLRLASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA
 PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS
 GETEYNPTEAR

N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 985-989

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697,
 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840,
 873-879, 912-918, 954-960

FIGURE 24

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC
CACGGCCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT
CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC
TGGACCCTTAACCTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT
CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC
GCACACTCCGTTACCACACTGGGTCAATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAG
ATAGCCCGGGTCATCTTGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC
CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC
TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA
AATGCGTTCATGCTACTCATGCGAAACATTGTGAGGGTGGTCGTCTGGACAAAGTCACAGA
CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCTGTCTTCTTTT
TTTTTCTCCGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC
TGGCTGCCCATCATGACCTCCATCCTGGGGGCTATGTCATCGCCAGCGGCTTCTTCAGCGT
TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG
GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC
GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAG**TGA**CAGCTCCGGCCCTGATCCAGGACTGC
ACCCACACCCACCGTCCAGCCATCCAACCTCACTTTCGCCTTACAGGTCTCCATTTTGTGGT
AAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG
AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC
GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGTTGCAGTGAGCCGAGA
TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA
AAAGATTTTATTAAAGATATTTTGTTAAGTCT

FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLGF
WTLNWWLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLLKILGKKN
EAPPDNKKRKK

FOSTP-TH026660

FIGURE 26

GAGTCTTGACCGCCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT
 GGCT**ATG**TTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC
 TTCTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTTC
 CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAACTGC
 ATTTCTTGAGCATAAAGAACAGTTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG
 ACCTATTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGG
 CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA
 TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT
 CAGGAAATGACAGTGATGGGTGAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA
 GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT
 CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG
 CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC
 CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTCTGCAGCG
 CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA
 CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC
 AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA
 GCGGCTCCAGGAGTTCCTTGACAGATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC
 AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA
 TTTGGGATGAAGGACATGCGCGTGCAGACTTTCAGCATTCATTTTGGGTTCAGCACAAAGTT
 TCTGGCCAGCGACGTGGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT
 CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG
 TACCATGGCCTGGAACTCGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC
 CTTTGACCAACCTCGTCATCTCCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC
 TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA
 AGTCCTTTGTGTGTTGACAAAGAACCGGCGCTGCAAACCTGCTGCCCCTGGTGATGGCTGCC
 CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC
 GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA
 TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT
 CTGGACGCACTTATTTCCCTCCTGTCC**TAG**GAAATTTGATTCTTCCAGAATGACCTTCTTATT
 TATGTAACCTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT
 TTTTATTAAATAAAATGCTTATTTTAGGAAA

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETA
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAIEDIFRDEEEDEEHSGNDS DGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGV LQRH
VSRHNHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR
LQEF LADMGLPLKQVKQKFQAMD ISLKENLREMI EESANKFGMKDMRVQTF SIHFGFKHKFL
ASDVVFATMSLMESPEKDGS GTDHF IQALDSLRSNLDKLYHGLELAKKQLRATQQT IASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKS FVCSTKNRRCKLLPLVMAAP
LSMEHGT VTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNF DLSVIELKAEDRSKFL
DALISLLS

099804.11501
TOSTT. T408660

FIGURE 28

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCGGTGGCTATGNTCGTGTCCGATTTCCGCA
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTTCTCTTCGTGGCCTCGGANGTGGAT
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT
GGTTCAGTTTCTGGGTGGCAAGAAGTGAAGTGCATTTCTTGAGCATAAAGAACAGTTTC
ATTATTTTATTCTCATAAAGTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA
CGATACCC

FIGURE 29

CAGGAACCCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC
 CACGAAGGGACGATACCAGAAAAACACCCCTCAACCCAAAGGAAATAGACTACAGCCCAATTG
 GCTGACTTTTGGCTATAGAAAAAAGAAAGGAACGAAAAGAGACAGTTTTTTTTTGGAAAGCTAA
 GTCTTCCCTTTATCGAGTCAAGAAACCCCCCTTCTTGAGCTATTTACAGCTTTTAACAATT
 GAGTAAAGTACGCTCCGGTCACC**ATG**GTGACAGCCGCCCTGGGTCCCCTCTGGGCAGCGCTC
 CTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC
 CAGCGGCTGCCAACGGTGTGTGACTCTGAGGACCCCTGGATCCTGCCCATGTATCCTCAG
 CCTCTTCCCTCCGGCCGCCCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC
 CTGAAGGGTGACAAAGGGGACCCAGGCCCAATGGGCCCTGCCAGGGTACATGGGCAGGGAGGG
 TCCCCAAGGGGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGGAGATGGGCAGCCCCG
 GCGCCCCGTGCCAGAAGCGCTTCTTCGCCTTCTCAGTGGGCCGCAAGACGGCCCTGCACAGC
 GGCGAGGACTTCCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGA
 CATGGCGACCGGCAGTTTGTCTGCCCTGCCCTGGCATCTACTTCTCAGCCTCAATGTGC
 ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTCTC
 CTGTACGCGCAGCCCAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGC
 CTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCA
 ACGACTTCGACACCTACATCACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGACT**TGAGGG**
 CCTCTGGGCCACCCCTCCCGCTGGAGAGCTCAGGTGCTGGTCCCCTGCCCTCAGGGCTCAG
 TTTGCACTGCTGTGAAGCAGGAAGGCCAGGGAGGTCCCCGGGGACCTGGCATTTCTGGGGAGA
 CCCTGCTTCTATCTTGGCTGCCATCATCCCTCCCAGCCTATTTCTGCTCCTCTCTTCTCTCT
 TGGACCTATTTTAAAGAGCTTGCTAACCTAAATATTCTAGAACTTTCCCAGCCTCGTAGCCC
 AGCACTTCTCAAACCTTGGAATGCATGCGAATCACCCGGGGTTCGTGTTAAATGCAGATTCT
 GACTCAGCAGGTCTGAGTGGGTCCAGGATTCTGTGTTTCTCATATGTTTCTGGGTGATGCTG
 ATGGGGTCAGTCTATGAACACACTGGAGCAACCAGGTTCTAGGACTTTCTCAATATTCTAG
 TACTTTCTGAACATTCTGGAATCCTCCCCACATTCTAGAATTCTCCCAACATTTTTTTTTTCT
 TGAGACAGAGTCTTGCTCTGTTGCCCAGGCTAGAGTGCAATCTCAGTTCACTGC
 AACCTCTGCCTCCCGGGTTCAAGCGATTCTTCTGCCTCAGCCTCCCTAGTGGCTGGGATTAC
 AGGCGCCTGCTACCATGCCTAGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTACCATA
 TTGGCCAGGCTGGTCTTGAACCTCTGACTTCAGGTGACCCACCCGCCCTCGGCCTCTCAAAT
 GCTGGGATTACAGGTGTGAGCCACCGTGCCTGGCCAATTCCAACATTCTTAAATTCTCTCAT
 CCCTCCAGGGCTCCCCGTGCTATGTTCTCTTTACCCCTTCCCCCTCTTCTCTTGCTCAGGCC
 TGCACCACTGCAGCCACCGTTTCAATTATTCAATTCAATAAACAAGTACTCACTCTGTGCT
 GGGTCCCCGGAAGGGTGAGGGGGTCAGACACAGGCCCTGCCCTGCCCTCAGTGACTGGCCA
 GTCCAGCCCAGGCGGGGAGAGATGTGTACATAGGTTTTTAAAGCAGACCCAGAGCTCATGGGG
 GCCTGTGTTCTGGGTGTTCAAGGTGCTGCTGGTCTCCATTACCCACTGCTCCCCAAGGCTGG
 TGGGACGGGGTCCCGGTGGCAGGGGCAGGTATCTCCTTCCCGTTCCCTCATCCACCTGCCAG
 TGCTCATCGTTACAGCAAACCCAGGGGGCCTTGCCAGGTCAAGGGTCTGTGAGGAGAGG
 ACCCAGGAGTGTGGGGGCATTTGGGGGGTGAAGTGCCCCCGAAGAATGGAACCCACACCCA
 TAGCTCTCCCCACAGCTGATACGGCATCCTGCGAGAAGACCTGCCCTCCTCACTGGGATCCC
 CTTCTGCTCCTCCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTTCCACCAAAGTCATCT
 GAACTTCCGTTTCCCCAGGGCCTCCAGCTGCCCTCAGACACTGATGTCTGTCCCAGGTGCT
 CTCTGCCCCCTCATGCCCCCTCTACCCGGCCAGTGCCCCGACTCTCCAGGCTTTATCAAGGTG
 CTAAGGCCCGGGTGGGCAGCTCCTCGTCTCAGAGCCCTCCTCCGGCCTGGTGCTGCCTTTAC
 AAACACCTGCAGGAGAAGGGCCACGGAAGCCCCAGGCTTTAGAGCCCTCAGCAGGTCTGGGG
 AGCTAGAGCAAAGGAGGGACCTCAGGCCTTCCGTTTCTTCTTCCAGGGTGGGGTGGCCTGGT
 GTTCCCCTAGCCTTCCAAACCCAGGTGGCCTGCCCTTCTCCCCAGAGGGAGGCGGCCCTCCGC
 CCATTGGTGCTCATGCAGACTCTGGGGCTGAGGTGCCCGGGGGGTGATCTCTGGTGCTCAC
 AGCCGAGGGAGCCGTGGCTCCATGGCCAGATGACGGAAACAGGGTCTGACCAAGTGCCAGGA
 AGACCTGTGCTATAAACACCCCTGCCTGATCCTGCCCTGCCTGACCCCGCCACGCCCTGCC
 GTCCAGCATGATTAAAGAATGCTGTCTCCTCTTGAAAAA

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT
FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

C1q domain proteins.

amino acids 144-178, 78-111 and 84-117

FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTG
 AGGCCGCGGCCTGCCCCGCCCCGGCTCCCTGCGCCGCGCCGCTCCCGGGACAGAAG**ATGTG**
 CTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG
 GCTGCCCATCCGGCTGCCAGTGCCAGCCAGCCACAGACAGTCTTCTGCACTGCCCCGCCAGGGG
 ACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT
 CACCATGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCAC
 AGAACCAGATCGCCAGCCTGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG
 GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGGCCT
 CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC
 TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC
 CTGCCCCGCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT
 CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACG
 AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG
 CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACAC
 CCGCATTGCCAGCTGCGGCCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG
 TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCGCTGCGG
 CTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCCTG
 GGTGCGCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA
 AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAGCCACCACC
 ACCACAGCCACAGTGCCACCACGAGGCCCGTGGTGCGGGAGCCACAGCCTTGTCTTCTAG
 CTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTCCA
 CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTC
 AATGGGGGCACATGCCACCTGGGGACACGGCACCACTGGCGTGCTTGTGCCCCGAAGGCTT
 CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCA
 CGCCGAGGCCACCACGGTCCCTGACCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC
 GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA
 TCGCAACCTATCGGGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTG
 AGTACACGGTCACCCAGCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGG
 CCCGGGCGGGTGCCGGAGGGCGAGGAGGCCTGCGGGGAGGGCCATACACCCCCAGCCGTCCA
 CTCCAACCACGCCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCG
 CCCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG
 GGGCGGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCT
 GGAAC TGAGGGGAGTGAAGGTCCCTTGAGGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAG
 AGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCAGGGCCTGGCCTC
 CAGTCACCCCCTCCACGCAAAGCCCTACATC**TAA**GCCAGAGAGAGACAGGGCAGCTGGGGCCG
 GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC
 CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCTCTGGA
 CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCCCTAACGTCCCCAGAAC
 CGAGTGCCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG
 GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCCACTCCAGGCGGA
 CCCTGGGGGCCAGTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG
 TGA CTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGCTTTA
 GGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCTG
 GGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTTGTAAGACAAACGATGATATG
 AAGGCCTTTTGTAAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 32

MCSRVP L L L L L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N
G I T M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R
R L E R L Y L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P
G I L D T A N V E A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G
N T R I A Q L R P E D L A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G
P W V R E S H V T L A S P E E T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S
S S L A P T W L S P T A P A T E A P S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E
G F T G L Y C E S Q M G Q G T R P S P T P V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L
T Y R N L S G P D K R L V T L R L P A S L A E Y T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A
V H S N H A P V T Q A R E G N L P L L I A P A L A A V L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G
P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E C E V P L M G F P G P G L Q S P L H A K P Y I

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT
CTTCATCATTATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACA**ATG**GAGACTCATCAG
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG
AAGAAAGGGAACCTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG
ACCCAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTTCAAG
TTTTCACTTCTGTCTCCAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGG
ATCTCAAAACCTTTGAATTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG
AAGAGTGTAACTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTTAAAGACTT
TGACACCATGCCATCTGTGAGGAAGCTGGCAACATGTCACACCTGGAAATCCTAGGTTTGA
GTGGGGCAAAAATACAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC
TTCTTAGGATTCAGAACTCTTCCTCATATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC
AAAAGTGCACATTTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTGCCTGATGGAATCA
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA
ATGCAACGAAATCTTAGTTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA
TTTACTCTGGGACGACCTTTTCTTATCTTACAATTTGTTTGGCATACATCAGTGGAACTACT
TTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTAC
TCAAATACTGTAATGAGAACTATAAAATGGAGCATGTACATTTTCAAGTGTGTTTACATTTCA
ACAGATAAAAATCTATTTGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATG
CACAAATGCCACACATGCTTTTCCCGAATTATCCTACGAAATTCCAATATTTAAATTTTGGC
AATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAAGTGCCTCACTTGAAAACCTCT
CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC
CCTTGGAACTCTGGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAATTTGCTCA
TGGCCAGAACTGTGGTCAATATGAATCTGTATACATAAATAAATTTGCTGATTCTGTCTCAG
GTGCTTGCCCCAAAAGTATTCAAATACTTGACCTAATAAATAACCAAATCCAAGTGTACCTA
AAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTTCTAAGTATG
CTCCCTGGATGCAGTCATTTTCAAGTACTTTTCAAGTCTGAACATGAAATGAAGTTCATTTCT
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAAGTCTAAATGCGGGAAGAA
ATCCATTTCCGGTGTACCTGTGAATTAATAAATTTTCAATCAGCTTGAAACATGATTCAGAGTC
ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAAGTAC
GTTAAAAGACGTTTCTATCTCCACGAATTATCTTGCAACACAGCTCTGTTGATTGTCAACATTG
TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCC
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA
AGAACAACCTCAAGAGAAATGTCCGATTCCACGCATTTTATTTTATACAGTGAACATGATTCTC
TGTGGGTGAAGAATGAATTGATCCCCAATCTAGAGAAGGAAGATGGTTCTATCTTGATTTGC
CTTTATGAAAGCTACTTTTGACCCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA
GAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAATGAGTGGTGCCATT
ATGAATTTCTACTTTGCCCCACCACAATCTCTCCATGAAAATTTCTGATCATATAATTTCTTATC
TTACTGGAACCCATTCCATTCTATTGCATTTCCACCAGGTATCATAAACTGAAAGCTCTCCT
GGAAAAAAGCATACTTGAATGGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGCAA
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAAGTGCAGACA
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT
ATAAAATCCCACAGTCTTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA
CAACCTTTTATGATGGCAATTTGACAATATTTATTTAAAATAAAAAATGGTTATTTCCCTTCATA
TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAAACACCTTCAAGTTTATAAGG
GCTTATGGAAAAGGTGTTTATCTCCAGGATTGTTTATAATCATGAAAATGTGGCCAGGTGC
AGTGGCTCACTCTTGTAATCCCAGCACTATGGGAGGCCAAGGTGGGTGACCCACGAGGTCAA
GAGATGGAGACCATCCTGGCCAACATGGTGAACCCCTGTCTCTACTAAAAATACAAAATTA
GCTGGGCGTGATGGTGACGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGT
GACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAAATGGAAAACATCC
TCATGGCCACAAAATAAGGTCTAATTCATAAATTATAGTACATTAATGTAATATAATATTA
CATGCCACTAAAAAGAATAAGGTAGCTGTATTTTCTGGTATGGAAAAACATATTAATAT
GTTATAAACTATTAGTTTGGTGCAAAACTAATTTGTTGGTTTTTGGCATTGAAATGCGATTGAA
ATAAAAGTGTAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGGGAGGTGGA
TTACAGGGAGCATTTGATTTCTATGTTGTGATTTCTATAATGTTTGAATTTGTTTAGAATGA
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTkMDIENLTISNAQMPHMLFPNYPTKFQYLNfANNILTDELfKRTIQLP
HLKTLILNGNKLETLsLVSCFANNTPLEHLdLSQNLLQHkNDENCsWPETVVMNLSYNKLS
DSVFRCLPKSIQILDlNNNqIQTVPKETiHlMALRELNIaFNfLTDLPGCSHfSRLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNfIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVGLAVAFcCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQlKRNVRfHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHhNLFHENSdHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKGLFWANLRAAINVNVLATREMYELQTfTELNEESRGSTISLM
RTDCL

FIGURE 35

GGGGGCTTTCTTGGGCTTGGCTGCTTGGAAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTGCGCGGGAAAGG
 GAGGGAAGAAGGAAGGGCGGGGCGGGCCCCCTGCGCCCGCCCCGCGCCTCTGCGCGCCCCCTGTCCGCCCCGGC
 CCAGCCCAGCCCAGCCCCGCGGGCCGGTCACACGCGCAGCCAGCCGCGCCGCTCCCGCGCCCCAAGCGCGCCGCT
 CTGCTGTGCCCTGCGCCCTTGCCCCGCGCCAGCTTCTGCGCCCGCAGCCCGCCCCGCGCCCCCGGTGACCGTGA
 CCCTGCCCCGGGCGCGGGGCGGAGCAGGC**CATG**TCCCGCCCCGGGGACCGCTACCCAGCGCTGGCCCTGGTGCTC
 CTGGCAGTGACCCTGGCCGGGTGCGAGCCCAGGGCGCAGCCCTCGAGGACCCTGATTATTACGGGCAGGAGAT
 CTGGAGCCGGGAGCCCTACTACGCGCGCCCGAGCCCGAGCTCGAGACCTTCTCTCCGCGCTGCTGCGGGGC
 CCGGGGAGGAGTGGGAGCGGCGCCCGCAGGAGCCCAGGCGGCCAAGAGGGGCCACCAAGCCCAAGAAAGCTCCC
 AAGAGGGAGAAGTCGGCTCCGGAGCCGCCCTCCACAGGTAAACACAGCAACAAAAAGTTATGAGAACCAAGAG
 CTCTGAGAAGGCTGCCAACGATGATCACAGTGTCCGTGTGGCCCGTGAAGATGTCAGAGAGAGTTGCCACCTC
 TTGGTCTGGAAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCA
 CATCGAGGGAGACTCAACATCCAGGCGGGCATTAATGAAAATGATTTTATGACGGAGCGTGGTGCGCGGGAAG
 AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATCACTGGTGTCATCACTCAAGGGA
 GGAACCTCCCTCTGGCTGAGTGACTGGGTGACATCCTATAAGGTGATGGTGAGCAATGACAGCCACACGTGGGTG
 ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTCTCAATGAGCT
 ACCCGTCCCCATGGTGGCCCGTACATCCGCATAAACCCCTCAGTCTCTGGTTTGATAATGGGAGCATCTGCATGA
 GAATGGAGATCCTGGGCTGCCACTGCCAGATCCTAATAATTAATATCACCGCCGGAACGAGATGACCACCACT
 GATGACCTGGATTTTAAGCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAATGAAATGTGTCC
 CAATATCACCGAATTTACAACATTTGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC
 ACCCTGGGGAGCATGAAGTCGGTGAGCCCGAGTTCCACTACATCGCGGGGGCCCCACGGCAATGAGGTGCTGGGC
 CGGGAGCTGCTGCTGCTGCTGGTGCACTTCGTGTGTCAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCT
 GGTGGAGGAGACGCGGATTACAGTCCCTCCCCCTCAACCCCGATGGCTACGAGAAGGCCTACGAAGGGGGCT
 CGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCCACGATGGAATGACATCAACAACAACCTTTCTGATTTA
 AACACGCTGCTCTGGGAGGCAGAGGATCGACAGAATGTCCCAGGAAAGTTCCCAATCACTATATTGCAATCCC
 TGAGTGGTTTCTGTGCGAAAAATGCCACGGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAAAAAATCC
 CTTTTGTGCTGGGCGGCAACCTGCAGGGCGGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGCGGTCCCC
 TGGAAGACGCAGGAACACACCCCCACCCCGATGACCAGTGTTCGCTGGCTGGCCTACTCCTATGCCTCCAC
 ACACCGCCTCATGACAGACGCCCCGGAGGAGGTGTGCCACAGGAGGACTTCCAGAAGGAGGAGGGCACTGTCA
 ATGGGGCCTCCTGGCACACCGTCTGCTGGAAGTCTGAACGATTTAGCTACCTTCATACAACTGCTTCGAACGT
 TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAAATACCGGGAATC
 TCTGATCGTGTTCATGGAGCAGGTTTCATCGTGGCATTAAAGGCTTGGTGAGAGATTACATGGAAAAGGAATCC
 CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCCTC
 CTGAACCTTGGAGAGTATGTGGTCACAGCAAAGGCCGAAGTTTCACTGCATCCACCAAGAATGTATGGTTGG
 CTATGACATGGGGGCCACAAGGTGTGACTTACACTTAGCAAACCAACATGGCCAGGATCCGAGAGATCATGG
 AGAAGTTTGGGAAGCAGCCCGTCAAGCTGCCAGCCAGGCGGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGT
 GGG**TGAC**CCCTCCTGGGCCCTTGAGACTCGTCTGGGACCATGCAAAATTAACCAACCTGGTAGTAGCTCCATAG
 TGGACTCACTCACTGTTGTTTTCTCTGTAATTCAAGAAGTGCTGGAAGAGAGGGTGCAATTGTGAGGCAGGTCC
 CAAAAGGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTCTTTGTTCCCATTTATCCAAATAACTTGGACAGAGCA
 GCAGAGAAAAGCTGATGGGAGTGAGAGAACTCAGCAAGCCAACCTGGGAATCAGAGAGAGAAGGAGAAGGAGGG
 GAGCCTGTCCGTTAGAGCCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCCTGTTTGCGTGCCAGCAAGG
 GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCCCGAGCTGGGCTGTCCCAATGTTACCA
 TTTGAGATGCTCCAGGCGTCTTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA
 ATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGACATCAGTGAGCCTCTTGAATCTGTTTAGTCTCCT
 TTTTCAACAAAGGAGTGTGTTTCAAGAAAGGAGAGAGAGGCTGAGATCATTAGGAGTTTGTGGGCAGCAAGCA
 TGGAGCTTCTTGACAAAATCTGGGTCCATAAACCAACCCCAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT
 CCCAGGTAGGGAGAGCCAGAGGTGCCAGCCTTCTGAAGGGCCAGAAAATTTAGCCTGGATCTCCTCTTTTAC
 CTGCTAGGACTGGAAGAGCCAGAAGTGGGGTGGCCTGAAGCCCTCTCTGCTTGAGGTATGCCCTGTGTG
 GAATTGAGTGCTCATGGGTTGGCCTCATATCAGCCTGGGAGTTATTTTATGATATGTAGAATGCCAGATCTTCCA
 GATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTATTGAATTAT
 CTTGCAAGAAAAAGTATGTCTCACTTTTTGTTAATGTTGCTGCCTCATTGACCTGGGAAAAATGAAAAAAA
 AATAAAGCAAATGGTAAGACCCTTAA

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYARPEPELETFSPPPLP
AGPGEWEERRPQEPRPPKRAKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPVMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGI DINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTLSTNMARIREIMEKFGKQPVSLPARR
LKLGRGRKRQRG

FIGURE 37

CTAAGAGGACAAGATGAGGCCCCGGCCTCTCATTTCTCCTAGCCCTTCTGTTCTTCTTGGCCAAAGCTGCAGGGG
 ATTTGGGGGATGTGGGACCTCCAATTTCCAGCCCCGGCTTCAGCTCTTTCCAGGTGTTGACTCCAGCTCCAGC
 TTCAGCTCCAGCTCCAGGTCCGGGCTCCAGCTCCAGCCGAGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT
 TTCCAATTTACCCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC
 CCGTGGACAGAGTGGAAACGCTTGAATTCACAGCTCATGTTCTTCTCAGAAAGTTTGAGAAAGAACTTTCTAAA
 GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAAACTGTTAAACCTAACTGTCCGAATTGACATCAT
 GGAGAAGGATACCATTTCTTACACTGAACTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAC
 TGGTCATACAGCTGAAGGAGAGTTTTGGTGAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT
 ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTGCGCGAGAAATCGTGGC
 TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTCGTCCACCCTCCTCCCACTC
 CAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTTT
 TCTTATCTATATGGTGCCTGGGGTAGGGATTACTCTCCCAGCATCCAAACAAAGGACTGTATTGGGTGGCGCC
 ATTGAAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTTGCTATTGTATATAA
 ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC
 ATGTACAACACCGGGAATATTGCCAGAGTTAACTGACCACCAACAGATTGCTGTGACTCAAACTCTCCCTAA
 TGCTGCCTATAATAACCGCTTTTTCATATGCTATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAATG
 GATTGTGGGTATTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAACTCAATGACACCACACTT
 CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGGTCT
 GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAAGAGG
 GCAAACCTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACCTATAACCCCTTTTGACCAGAAA
 CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTCTGTCTTGCAGAAGCCCCAGTAAGCTGTTTA
 GGAGTTAGGGTGAAAGAGAAAAATGTTTGTGAAAAAATAGTCTTCTCCACTTACTTAGATATCTGCAGGGGTGT
 CTAAAAGTGTGTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT
 TGATTTGGTGAGTTCTCTTGGGAATCATCTGCCTCTTCAGGCGCATTTTGCAATAAAGTCTGTCTAGGGTGGGA
 TTGTGAGAGGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA
 GGAATTAAGGAACCTTAAACCTCAGTATGGCGTCTAGGGATTCTTGTACAGGAAATATTGCCAATGACTAGTC
 CTCATCCATGTAGCACCCTAATTCTTCATGCCTGGAAGAAACCTGGGGACTTAGTTAGGTAGATTAATATCT
 GGAGCTCCTCGAGGGACCAATCTCCAACTTTTTTTTTCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG
 CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTT
 ATGCATTAAATTTGACATGGCAAATAAATCCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTCTCTC
 ATTGTCACCTTACTAAAAGTCAGTAGAATCTTCTACCTCATAACTTCCTTCCAAAGGCAGCTCAGAAGATTAG
 AACCAGACTTACTAACCAATTCACCCCCCACCACCCCCCTTCTACTGCCTACTTTAAAAAAATTAATAGTTTT
 CTATGGAACCTGATCTAAGATTAGAAAAATTAATTTTCTTTAATTTTCAATTATGGACTTTTATTTACATGACTCTA
 AGACTATAAGAAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA
 TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTGCCTTTGTTTAAGCCTGGAACCTGTAAGAAAAT
 GAAAATTTAATTTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT
 TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT
 TCCTTTGATGTTCAAGTCCTAGTCTATAGGATTGGCAGTTTAAATGCTTTACTCCCCCTTTTAAAAATAAATGAT
 TAAATGTGCTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 38

MRPGLSFLLALLFFLGQAAGDLGDVGPPIPSPGFSSFPGVDSFFFSSSSSRGSSSSSRSLGS
GGSVSQLFSNFTGSVDDRGTCCQCSVSLPDTTFPVDRVERLEFTAHVLSQKFEKELSKVREYV
QLISVYEKKLLNLTVRIDIMEKDTISYTELDFELIKVEVKEMEKLVIQLKESFGGSSEIVDQ
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKTKLKECEASKDQNTPVVHPPPTPGSCGH
GGVNVISKPSVVQLNWRGFSYLYGAWGRDYSPOHPNKGLYWVAPLNTDGRLLLEYRLYNTLD
DLLLLYNARELRITYGQSGTAVYNNNMYVNMVNTGNIARVNLTNTIAVTQTLPNAAYNRR
FSYANVAWQDIDFAVDENGLWVIYSTEASTGNMVISKLNDTTLQVLNTWYTKQYKPSASNAF
MVCGLYATRMTNTRTEEIFYYYDTNTGKEGKLDIVMHKMQEKVQSINYNPFDQKLYVYNDG
YLLNYDLSVLQKPQ

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCACTCCAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA
TGTACAACACCGGGNATATTGCCAGAGTTAACCTGACC

FIGURE 40

TCTCGCAGATAGTAAATAATCTCGGAAAGGCGAGAAAGAAGCTGTCTCCATCTTGTCTGTAT
 CCGCTGCTCTTGTGACGTTGTGGAG**ATG**GGGGAGCGTCCTGGGGCTGTGCTCCATGGCGAGCT
 GGATACCATGTTTTGTGTGGAAGTGCCCCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAC
 AACTCCACTGTAAC TAGATTGATCTATGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTG
 TGTAATGTTGATACCAGGAATGGAAGAACAACGAATAAGATTCCCTGGATTTTGTGAGAATG
 AGAAAGGTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTT
 GGTTTGGCTATGTTCTATCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGA
 TCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTA
 TTATTGGGGCATTCTTCATTCCAGAAGGAAC TTTTACAACGTGTGTGGTTTTATGTAGGCATG
 GCAGGTGCCTTTTGTTCATCCTCATACAAC TAGTCTTACTTATTGATTTTGCACATTCATG
 GAATGAATCGTGGGTGAAAAAATGGAAGAAGGGAACTCGAGATGTTGGTATGCAGCCTTGT
 TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCCTGTTCTTTGTCTAC
 TACACTCATCCAGCAGTTGTTTCAGAAAACAAGGCGTTCATCAGTGTCAACATGCTCCTCTG
 CGTTGGTGCTTCTGTAATGTCTATCTATCGCCAAAAATCCAAGAATCACAACCAAGATCTGGTT
 TGTTACAGTCTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT
 GAACCAGAAAACAATGCAACCCAAGTCTACTAAGCATAATTGGCTACAATACAACAAGCAC
 TGTCCCAAAGGAAGGGCAGTCACTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTCT
 TCTTTTGTGTGTGTATTTTATTTCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAA
 CTGACTCTAACAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC
 ACTGGAGGATGGGGACGATGTTCAACGAGCTGTAGATAATGAAAGGGATGGTGTCACTTACA
 GTTATTCCTTCTTTCACCTTCATGCTTTTCTGGCTTCACCTTTATATCATGATGACCTTACC
 AACTGGTCCAGGTATGAACCTCTCGTGAGATGAAAAGTCAGTGGACAGCTGTCTGGGTGAA
 AATCTCTCCAGTTGGATTGGCATCGTGTGTATGTTTGGACACTCGTGGCACCACCTTGTTC
 TTACAAATCGTGATTTTGGAG**TGAG**TGAGACTTCTAGCATGAAAGTCCCACTTTGATTATTGC
 TTATTTGAAAACAGTATTTCCCAACTTTTGTAAAGTTGTGTATGTTTTTGCTTCCCATGTAAC
 TTCTCCAGTGTTCTGGCATGAATTAGATTTTACTGCTTGTCATTTTGTTATTTTCTTACCAA
 GTGCATTGATATGTGAAGTAGAATGAATTGCAGAGGAAAGTTTATGAATATGGTGATGAGT
 TAGTAAAAGTGGCCATTATTGGGCTTATCTCTGCTCTATAGTTGTGAAATGAAGAGTAAAA
 ACAATTTTGTGTGACTATTTTAAATATATAGACCTTAAGCTGTTTTAGCAAGCATTA
 GCAAATGTATGGCTGCCTTTTGAATATTTGATGTGTGCTTGGCAGGATACTGCAAAGAAC
 ATGGTTTTATTTTAAATTTTATAACAAGTCACTTAAATGCCAGTTGTCTGAAAAATCTTATA
 AGGTTTTACCCTTGATACGGAATTTACACAGGTAGGGAGTGTTTAGTGGACAATAGTGTAGG
 TTATGGATGGAGGTGTCGGTACTAAATTGAATAACGAGTAAATAATCTTACTTGGGTAGAGA
 TGGCCTTTTGCCAACAAAGTGAAGTGTGTTTGGTGTGTTTTAAACTCATGAAGTATGGGTTCA
 GGAAATGTTTGGAACTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTAGA
 AGGAAGTGTGTTTGAAGTCACTTTGAAAGTTAGTTTTGGGCCAGCACGGTAGCTCACCTTT
 GGTAATCCCAGCACTTTGGGAGCTTAAGTGGGTAGATTACTTGAGCCCAGGAATTCAGACCA
 GCTTGGCACATGGTGAACCTGTTCTATAAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTC
 CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCAGAGCCAAAGGTTGCAGTGAGCAAGTCA
 CGTCACTGCACTCTAGCTGGCACAGAGTAAGCCAAAAAATATATATATATTGAAATCAAGG
 AGGCAAAATTTTGAAGGGAAGGAAGTAACTGCAAAACCACTAGGCTTTAGTAGGTACTTAT
 ATAAAATCTAGTCCAGTTCTCTCATT TAAAAAATGAAGACACTGAAATACAGACTTAAATA
 GCTCAGATAGCTAATTAGGAAATTTCAAGTTGGCCAATAATAGCATTCTCTGACATTTAA
 AAATAATTTCTATTCAAATAACATGCATATTGATTTACACCTCATACTGTGATAATTAAATGT
 GATGTGGATTGCTGGTGTCCAGCATGACCATAAACAGGTCAGAAGAATGATGGAATGTTTT
 AGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAGATGTTTAAATGCTTTTGTAT
 TTACTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGT
 ATGAGAGTTTAGTTATTTGTATGTGTCACTAGTCTAATGAAGCTTTTAAATCTACAATT
 TCTTCTTTAAAAATATTTATTAATGTGAATGGAATATAACAATTCAGCTTAATCCCCAACCC
 TTATTCTGTGTGTAGACATTGTATTCCACAATTTTGAATGGCTGTGTTTTACCTCTAAATAA
 ATGAATTCAGAGAAAAA

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMYLTSAMTNEPETNCNPSSLIIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVTSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVWKISSWIGI
VLYVWTLVAPLVLNDRFD

FIGURE 42

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAACAANTCCACTGTAAGTAGATTGATCTA
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT
GTTGATACCAGGAATGGAAGAACAACCTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG
GGGC

FIGURE 44

AAGAAGCTGTCTCCATCTTGTCTGTATCCGCTGCTCTTGTGAACGTTNTGGAGATGGGGAGC
GTCCTTGGGGTTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCCCCGTGTT
TGCTATGCCGATGCTGTCCTAGTGGAAACAACCTCCACTGTAAC TAGATTGATCTATGCACTT
TTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAGAACAAC
GAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATTTTGTTG
GCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTA
CTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGTT
CTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

FIGURE 45

GCTGTCCTTAGTGGAAACAANTCCAACCTTGTAACCTGGATTGATCTATGCACTTTTTCTTG
CTTGTTGGAGTATGTGTAGCTTTGTGTAATGTTGTTCCCAGGATTGGANGAACAACTGAATA
AGATTCCTGGATTTTTGTGAGAATGAGAAAGGTGTTGTCCCCCTGTAAACATTTTTGGTTGGC
TATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTCTCTTTACT
AATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCT
TTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGAACTTTT
ACAACTGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTGTTTCATCCTCATACAAC TAGT
CTTACTTATTGATTTTGCACATTCATGGAATGAATCGTGGGTGAAAAAATGGAAGAAGGGA
ACTCGAGATGTTGGTATGCAGCCTTGTTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTA
GTTGCTATCGTCCTGTTCTTTGTCTACTACACTCATCCAGCCAGTTGTTTCAGAAAACAAGGC
GTTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAATG

FIGURE 46

CTCGGGCGCGCACAGGCAGCTCGGTTTGCCCTGCCATTGAGCTGCGGGTCGCGGCCGGCGCCGGCCTCTCCAAT
GGCAAATGTGTGTGGCTGGAGGCGAGCGGAGGCTTTCGGCAAAGGCAGTCGAGTGTTGCAGACCGGGGCGAG
TCCTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCGGGCGGGCTGTGCG
ACTCCCCGCGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTAC
GTCGTTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCAGGATATTTGATTTATTAGCGATGCC
CCCTGGTTTTGTGTGTGTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCC
TGGGCGAATCCACATCTGTTTCAACTCTCCGCGGAGGGCGAGCAGGAGCGAGAGTGTGTGGAATCTGCGAGTG
AAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAAGAACACCAGAT
CAGCAAAAAAAGAAGATGGGCCCCCGAGCCTCGTGTGTGCTGTGCTGTCCGCAACTGTGTTCTCCCTGCTGGG
TGGAAGCTCGGCCCTTCTGTGCGACACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCA
ACATCATCTTGGTGTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC
ATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCTCACGCTCCTC
CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAAGTGTCTCGCCCTCCTGGC
AGGCACAGCAGAGAGCCGACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTCTTCGGGAAG
TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGTCCGACTCCTTAAAAACTCCCG
CTTTTATAACTACACGCTGTGTCGGAACGGGGTGAAAGAGAAGACACGGCTCCGACTACTCCAAGGATTACCTCA
CAGACCTCATCACCATGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCTCTC
ATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTCAGCCCCACAATATTCACGCCTCTTCCAAACGC
ATCTCAGCACATCAGCCGAGCTACAACCTACGCGCCCAACCCGGACAAACACTGGATCATGCGTACACGGGGC
CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGCCCTCATGTGCGGTGGAC
GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC
CGACCACGGTTACACATCGGCCAGTTTGGCCTGTTGAAAGGGAATCCATGCCATATGAGTTTGACATCAGGG
TCCCGTTCTACGTGAGGGGCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCTCAACATTGACCTG
GCCCCACCATCCTGGACATTGCAGGCCTGGACATACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCT
GGACACGGAGCGGCCGTTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG
AGAGAGGCAAGCTGTACACAAGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAAGTTTCTGCCCAAGTAC
CAGCGTGTGAAGGACCTGTGTACGCTGTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG
TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC
TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCTGCACCTGTGACAGCGGGGACTACAAGCTC
AGCCTGGCCGGACGCCGAAAAAACTCTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCG
CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA
CCAAGCGGCACTGGCCAGGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGC
CTTCCCGACTACTCAGCCGCCAACCCCATTAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA
GTGTACCTTGGACCTGTACAAGTCCCTGCGAGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAA
CCCTGCAGAACAAAATTAAGAACCTGAGGGAAGTCCGAGGTACCTGAAGAAAAAGCGGCCAGAAGAATGTGAC
TGTCACAAAATCAGTACCACACCCAGCACAAAGGCCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTTCAG
GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGC
TCAAGCGCCTGCAGAACACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACCAGCACTGG
CAGACGGCGCCTTCTGGACACTGGGGCCTTCTGTGCTGCACCAGCGCCAACAATAACACGTACTGGTGCAT
GAGGACCATCAATGAGACTCACAATTTCTCTTCTGTGAATTTGCAACTGGCTTCTAGAGTACTTTGATCTCA
ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCTCAACCAGCTACACGTACAG
CTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGAACCCCGGACTCGAAACATGGACCTGGATGGAGG
AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG
GACAACTGTGGGAAGGCTGGGAAGGTTAAGAAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA
CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC
AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCC
ATTTTTGCCCCGTGTTTTGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCGTATCAAAAAGTC
ACCACTAACCTTCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCGAGAGAGATTTCTTGGAATTTTC
TCCCAAGGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAAAGCAGTCTGTCTAAATCCTCTTATCTT
TTGGTTTGTACAAAAGAAGGAACATAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG
TTTGACAAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCTTGGTTGCCTCTGAAGAAA
CTGCTTCATTGTATATATGTGACTATTTACATTAACATCGGAACCTTTTAGGGGAACCTAATAAGAAAT
CCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAGAAAAA

FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTTMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHE
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQVRKDLQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSGDYKLSLAGRRKKLFFKKKYKASYVRSRSIRSVAI EVDGRVYHVGLGDAAQ
PRNLTKRHWPGAPEDQDDKDGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL
HPFRKGLQEKDKVWLLREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWCMRTINETHNFLCFEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

FIGURE 48

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA
GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAGATGCTGGGCCTCCTGGGGAGCACAGCCC
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCTGCTGCTGCTGCTGCTGCTGGCCACC
TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGAAA
CCGAGTCCGCCGGGCCCAGCCTTGGCCCTTCCGGCGGGGGCCACCTGGGAATCTTTCACC
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA
CGCTCGCTGAGGCTGCTGTCGCCGGTGCCTGTGGACAGCAGCTGCCCCTGCCCTCCCATCTG
TCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG
GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG
GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGACTCCCAGTGAGCCCCAGAAATG
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

49/330

FIGURE 49

MLGLLGSTALVGWITGA AVALLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRR AQPWPFR
RRGHLGIFHHHRHPGHVSHV PNVGLHHHHHPRHTPHHLHHHHHPRHHHPRHAR

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 50

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCC**ATG**TCGGACCTGCTA
CTACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG
GTACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTACCCCCCATCCGCAACGTCA
CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC
TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC
CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC
CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC
CATGTGGTGACAGCCACCTTCCCCTACACCACCATTCTGTCCATCTGGCTGGCTACCCGCCG
TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG
AGATCTACCAGGAAGACCAGATCCATTTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTAT
GTGCCTGAGATGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA
GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC
CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTACCTGGGGCGAGCAGCCGTGGCTGGGAT
GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA
GGAGCTGGACTTGAGGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC
CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAG**TAA**CCC
ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT
CCAGCCCTCTTCCTCCTTCCTCTGGGGGAGGAGGGGTTCCTGAGGGACCTGACTTCCCCTGC
TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA
GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTCAGACTC
ACAGTGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAA
AAAAA

FIGURE 51

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRAVGSILSEGEESPSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATRRVHPALDITYIKERKLCAYPRLEIYQEDQIHMCPLAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSVSVLEVSPGSRETSAAATLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

FIGURE 52

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCTTGGTGCCCCGGCTCCCT
GCCCCGCGCCCAGTC**ATG**ACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA
CCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGA
GACACGCTTCACATACACTACCGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCCTTCTCACTTGGCCTAT
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT
GATTGCACTAATCCGAGCCAACCTACTGGCTAAAGCTGGTGAAGGGCATTCTTGCCTCTGGTAG
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAAT
AGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAA**TA**
ATAAATAATAAATTTTAAAAAACTTAAAAAAAAAAAAAAAAAAAA

FIGURE 53

MTLRPSLLPLHLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGLTLHI
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS
KKKLKEEKRNKSKKK

090904.450
T05T"YH03060

FIGURE 54

CCCGGGAACGTGTTCTGGCTGCCGCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCTGC
CCCGCGCCCAGTCATGACCCTGCGCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCTGC
TGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGACC
CTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGAGA
CACGCTTCACATACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCTGA
CCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGAGT
CTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCCTTCTCACTTGGCCTATGG
AAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCTGA
TTGCACTAATCCGAGCCAACACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAGGG
ATGGCCATGGTGCCACCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAATAGA
CCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATAATA
AATAATAAATTTTAAAAA ACTTA

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAA
CAAGAGCAAAAAGAAATAATAATAATAAATTTTAAAAAACTTAAAA

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG
 TGAGGCGGGCCGGCGCGGCGGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG
 ACCTGAAAAAAATGCTCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG
 GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT
 TATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCT
 GTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA
 GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG
 TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG
 CTAAAGAAAAAGACATAGTATAACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT
 TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAGTGAACACATCTGAT
 TTCCACAGCACAAACAGCCCTGCATGGGTTTGTGTTGTTTTTTTACTGCTCACTCCCAACCTT
 TTGTAATGCCATTTTCTAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT
 AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACCT
 ATAAAATGTTAGAGGAACTTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT
 TGTA AAAAATAAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA
 TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT
 CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTAGTTTTTAAATATTCCGTGG
 TCAAAATTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT
 GTA ACTGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT
 GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTCCGCTGTGCCTCTCATT
 CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC
 CACATCCACCACTG

57/330

FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWIIIDA AVIYPTMKDFNHSYHACGVI
ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILEGGYVAKEK
DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 58

TTCTTGGCTAAAATCGGGGGAGTGAGGCGGGCCGGCGCGGCGGACACCGGGCTCCGGAACC
ACTGCACGACGGGGCTGGACTGACCTGAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATG
CTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTAC
TATTTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGAT
TTCAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGCCTTCCTAATGATTAATGC
AGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTG
CTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGG
ATTCTTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATT
TTTCCAGAATGCCTTCATCTTTTTTTGGAGGGCTGGTTTTTTAAGTTTGGC

TGGACGGACCTGAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATG TTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC
ATCTTTTTTTGGAGGGCTGGTTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTTCCAGAATGCC

FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTGTGGCTTTTNGTTGGTTTCATGTTGGCCTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

FIGURE 62

GGGAGGCTGTGNCCGTTTTGTTTTNTTGGCTAAAATCGGGGGAGTGAGGCGGCCCGGCGCGG
CGNGACACCGGGTTCCGGGAACCATTCACGACGGGGTGGACTGACCTGAAAAAATGTTTG
GATTTNTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATT
GCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGT
TATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCA
TAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAA
GGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGG
ATNTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAG
TATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATNTTTTTTTGGAGGGCTG

FIGURE 63

CGACGCCGGCGT**GATG**TGGCTTCCGCTGGTGTCTGCTCCTGGCTGTGCTGCTGCTGGCCGTCC
TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC
AAACGGCCCCCAGCGCCCTGGTAAGTACAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC
TTTTTCAGCCAACCAAGTGCCGGAGAAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG
GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCCTGGTGTGGAACAACAT
ACCAAGGCAGGGGGCTGCTGTACATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAAT
CCATTACATTGGGCGTATGGAAGAGGGGCAGCATTGGCCGTTTTATCTTGGACCAGATCACTG
AAGGGCAGCTGGACTGGGCTCCCCCTGTCTCTCCTTTTGACATCATGGTACTGGAAGGGCCC
AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAGCCTACATTAGGGCCTCAAGGA
GAAGTTTCCACAGGAGGAAGCTATCATTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA
GTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCCTCCCATTGCCCCGTGGTTCAGCTCCTCGAC
AGGTGTGGGCTGCTGACTCGTTTCTCTCCATTCTTCAAGCATCCACCCAGAGCCTGGCTGA
GGTCTCTGCAGCAGCTGGGGGCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCA
CTTACGGTGTACCCCCAACACAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC
ATGAAAGGAGGCTTTTATCCCCGAGGGGTTCCAGTGAAATTGCCTTCCACACCATCCCTGT
GATTAGCGGGCTGGGGGCGCTGTCTCACAAAGGCCACTGTGCAGAGTGTGTGCTGGACT
CAGCTGGGAAAGCCTGTGGTGTGAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTATTGC
CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGC
CCGCTGCCTGCCAGGTGTGAAGCAGCAACTGGGACGGTGCGGCCCGGCTTAGGCATGACCT
CTGTTTTATCTGCTGCGAGGCACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTAT
GTTTACTATGACACGGACATGGACCAGGCGATGGAGCGCTACGTCTCCATGCCCAGGGGAAGA
GGCTGCGGAACACATCCCTCTTCTCTTCTCGCTTTCCCATCAGCCAAAGATCCGACCTGGG
AGGACCGATTCCCAGGCCGCTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTTT
GAGGAGTGGCAGGCGAGCTGAAGGGAAAGCGGGGCAGTACTATGAGACCTTCAAAAACCTC
CTTTGTGGAAGCCTCTATGTGAGTGGTCTGAAACTGTTCCACAGCTGGAGGGGAAGGTGG
AGAGTGTGACTGCAGGATCCCCACTCACCAACCAGTTCTATCTGGCTGCTCCCCGAGGTGCC
TGCTACGGGGCTGACCATGACCTGGGCCGCTGCACCCCTTGTGTGATGGCCTCCTTGAGGGC
CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTACCTGTGGACTGGTCG
GGGCCCTGCAAGGTGCCCTGCTGTGCAAGCGCCATCCTGAAGCGGAACCTTGTACTCAGAC
CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAAT**TAG**TTCATCAGGGAGG
AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
CATTAGTTTCCTTGACAGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG
TTTAAATCACAAATCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
TCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATAGCTTGATG
TCTCATGACGAGCGGCGCTCTGCATCCCTCACCCATGCCTCCTAACTCAGTGATCAAAGCGA
ATATTCCATCTGTGGATAGAACCCCTGGCAGTGTGTGTCAGCTCAACCTGGTGGGTTTCAAGTTC
TGTCCTGAGGCTTCTGCTCTCATTCAATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGG
GAAAAGGGAGACTAATGAGGCTTAACTCAAAACCTGGGCGTGGTTTTGGTTGCCATTCCATA
GGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTTCAAGTGGCTCTTCAAGGGACAGGAAAT
GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGTTAGTA
GGGTGCATGTGAGATGATCATATCCAATTCATATGGAAGTCCCGGGTCTGTCTTCTTATCA
TCGGGGTGGCAGCTGGTTCTCAATGTGCCAGCAGGGACTCAGTACCTGAGCCTCAATCAAGC
CTTATCCACCAATAACACAGGGAAGGTGATGCAGGGAAGGTGACATCAGGAGTCAGGGCA
TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGG
CACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGAAGTCACATCAGAAAAGGGA
AAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTTGCAAGTTAATTAGCACATGTGAGGG
TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAAGTTATGTCTTTG
GTATCAGACATACGAAAGGTCTCTTTGTAGTTCTGTGTTAATGTAACATTAATAAATTTATTG
ATTCCATTGCTTTAAAAAAAAAAAAAAAAA

FIGURE 64

MWLPLVLLLLAVLLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKARKKVLKQAFSAN
QVPEKLDV VVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG
RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ
EEAII DKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ
LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA
GGAVLTKATVQSVLLDSAGKACGVS VKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP
GVKQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTMDQAMERYVSMPREEEAAEH
IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEWQAELKGKRGSDYETFKN SFVEA
SMSVVLKLFQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI
PNLYLTGQDIFTTCGLVGALQGALLCSSAILKRNLYS DLKNLDSRI RAQKKKN

FIGURE 65

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCTA
 GGGGTTGGCACCGGCCCGAGAGGAGG**ATG**CGGGTCCGGATAGGGCTGACGCTGCTGCTGTG
 TGGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT
 TAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTAAAGGACCATACTACTGCAGGCAGA
 GTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGTCTGAATTAGAATCCTCTATTCAAGA
 AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAGAAGATATCAGCTTTCTAG
 AGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTACGGAAACCAGCTTTG
 ACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCTTTTCTTAGATAA
 GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTGCTACAACCT
 ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCTAAGAGA
 CGGCAGATGCAGGAAGCAGAAATGATGTATCAAAGTGAATGAAAATCCTTAATGGAAGCAA
 TAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
 CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC
 CAGGCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC
 TCTTGGCTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTAT
 ATTATACATTTGGAGCTCTTGGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTT
TAGTGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTT
 TCAGCTTTCATGATCCAGATTTGCTTGTATTAAGACCAAATATTCAGTTGAACTTCCTTCAA
 ATTCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA
 ATTTTTCTTTAAAATGATTAGTTTGGCTGATTGCCCCTAAAAAGAGAGATCTGATAAATGGC
 TCTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTACATTAGATTATCATAA
 TTTTAAAAATTTTCTTTAGTTTTTCAAATTTTGTAAATGGTGGCTATAGAAAAACAACAT
 GAAATATTATACAATATTTTGCAACAATGCCCTAAGAATTGTTAAAATTCATGGAGTTATTT
 GTGCAGAATGACTCCAGAGAGCTCTACTTTCTGTTTTTTACTTTTCATGATTGGCTGTCTTC
 CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC
 CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACCTATTTTAAATAAA
 ATTATGTCTAAGATTAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 66

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLDSKTTLTSDSVKDHTTAGRVVAGQIFLD
SEESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEPPKKVRKPALTAIEGTAHG
EPCHFPPFLFDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRQMQEAMM
YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFEEK
LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

CTTCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAAGAACTTCTACTTGCCT
GCCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG
CTCAACTCTCCTGCACGCTCAGCCCCCAGCACGTACCATCAGGGACTACGGTGTGTCTCTGG
TACCAGCAGCGGGCAGGCAGTGCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCAACAATGCCT
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC
TACGGCTTTAGTCCC**TAG**GGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTTCCT
GCCCCTGACCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGG
TTAATAATATTCAACATGTCAACAAC

FIGURE 68

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPQVAQLSCTLSPQHVTIRDYGVSWYQQRAG
SAPRYLLYYRSEEDHHRPADIPDRFSAKDEAHNACVLTISPVPEDDADYYCSVGYGFSF

09904133
T03T1 T03060

CGCCGCCCGCCCCGAGACCGGGCCCGGGGGCGCGGGGCGCGGGGATGCGGCGCCCGGGGGCGG
CGATGACCGCGGAGCGCACGCCGCGGGCCCGGCCCTGACCCCGCCGCCCGCCCGCTGAGCCC
CCCGCCGAGGTCGGGACAGGCCGAGATGACGCGCCGAGCCCCCTGTTGCTGCTCCTGCTGCCGC
CGCTGCTGCTGGGGCCCTTCCCACCGGCCCGCCCGCCGAGGCCCCCCAAAGATGCGGGAC
AAGGTGGTCCCACGGCAGGTGGCCCGGCTGGGCGCACTGTGCGGCTGCAGTGCCAGTGGCAGTGA
GGGGGACCCGCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA
GCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC
GTGTACGTGTGTCAAGGCCACCAACGGCTTCGGCAGCCCTGAGCGTCAACTACACCCCTCGTCGT
GCTGGATGACATTAGCCCCAGGGAAGGAGAGCCTGGGGCCCGACAGCTCCTCTGGGGGTCAAG
AGGACCCCGCCAGCCAGCATGTGGGCACGACCGCGCTTTCACACAGCCCTCCAAGATGAGGCGC
CGGGTGATCGCACGGCCCGTGGGTAGCTCCGTGCGGCTCAAGTGCCTGGCCAGCGGGCACCC
TCGGCCCCGACATCACGTGGATGAAGGACGACCAGGCCCTTGACGCGCCAGAGGGCCGCTGAGC
CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCCTGCGGCCGGAGGACAGCGGCAATAAC
ACCTGCCGCGTGTGAACCGCGCGGGCGCCATCAACGCCACCTACAAGGTGGATGTGATCCA
CGGGAACCCGTTCGAAGCCCGTGCTCACAGGCACGACCCCGTGAACACGACGCGTGGACTTCG
GGGGGACCACGTCCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG
AAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACCTCCACCATCGATGTGGGCGGCCAGAA
GTTTGTGGTGCTGCCACGGGTGACGTGTGTCGCGGCCCGACGGCTCCTACCTCAATAAGC
TGCTCATCACCCGTGCCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGGCGCCAACAC
ATGGGCTACAGCTTCCGACAGCGCCTTCTCACCGTGCTGCCAGACCCAAAACCGCCAGGGCC
ACCTGTGGCCTCCTCGTCTCGGCCACTAGCCTGCCGTGGCCCGTGGTTCATCGGCATCCCAG
CCGGCGCTGTCTTCATCCTGGGCACCCTGCTCCTGTGGCTTTGCCAGGCCCAGAAGAAGCCG
TGCACCCCGCGCCTGCCCTCCCCCTGCCCTGGGCACCGCCCGCCGGGGACGGCCCGCGACCG
CAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCGCTCAGCGCTGGCCCTGGTGTGGGGCTGT
GTGAGGAGCATGGGTCTCCGGCAGCCCCAGCACTTACTGGGCCAGGCCAGTTGCTGCTGGC
CCTAAGTTGTACCCCAAACCTCTACACAGACATCCACACACACACACACACACACTCTCACAC
ACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGCTAGACGGCACCGT
ATCTGCAGTGGGCACGGGGGGGGCCGGCCAGACAGGCAGACTGGGAAGATGGAGGACGGAGCT
GCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCAGGCAGTCTGTGTG
TGAGGCATAGCCCCCTGGACACACACACACACAGACACACACTACCTGGATGCATGTATGCAC
ACACATGCGCGCACACGTGCTCCCTGAAGGCACACGTACGCACACGCACATGCACAGATATG
CCGCTGGGCACACAGATAAGCTGCCCAAATGCACGCACACGCACAGAGACATGCCAGAACA
TACAAGGACATGCTGCCCTGAACATACACACGCACACCCATGCGCAGATGTGCTGCCCTGGACA
CACACACACACAGGATATGCTGTCTGGACGCACACAGTGCAGATATGGTATCCGGACACA
CACGTGCACAGATATGCTGCCCTGGACACACAGATAATGCTGCCCTGCACACACATGCACCG
ATATTGCCCTGGACACACACACACACACACACGCGTGCACAGATATGCTGTCTGGACACGCACAC
ACATGCAGATATGCTGCCCTGGACACACACTTCCAGACACACGTGCACAGGCGCAGATATGCT
GCCCTGGACACACGCAGATATGCTGTCTAGTCACACACACACGCAGACATGCTGTCCGGACAC
ACACACGCATGCACAGATATGCTGTCCGACACACACACGCACGCAGATATGCTGCCCTGGAC
ACACACACAGATAATGCTGCCCTCAACACTACACACAGTGCAGATATTGCCCTGGACACACACA
TGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACG
CACGCACACATGCAGATATGCTGCCCTGGGCACACACTTCCGGACACACATGCACACACAGGT
GCAGATATGCTGCCCTGGACACACACAGATAATGCTGCCCTCAACACTACACACAGTGCAGA
TATTGCCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACAGTGCAGATA
TGCTGTCCGGATACACACACGCACGCACACATGCAGATATGCTGCCCTGGGCACACACTTCCGGA
CACACATGCACACACAGGTGCAGATATGCTGCCCTGGACACACGCAGACTGACGTGCTTTTGG
GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT
CCCTGCTCCACCGTCACTCCCCCAACTCTGCCCGCCTCTGTCCCCCGCTCAGTCCCCCGCCTC
CATCCCCGCTCTGTCCCCCTGGCCTTGGCGGCTATTTTGGCCAGCTGCCCTGGGTGCCCAGG
AGTCCCCCTACTGCTGTGGGCTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAG
CCCATGGCTAGTGGCTCATCCCCAGTGCAATTCTCCCCCTGACACAGAGAAGGGGCCCTGGTA
TTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTGGGTTGCAGGGAC
TGTGGTCTCTCCTAGGGGCCCGGACCCCGCCTGGTCTTTGAGCCATGCTGATGCCACAACCCC
GTCCAGGCGTACACAGCCCCCCCCACCCCACTGTGCTGGTGCCCGCAGATCTCTGTAATTTTA
TGTAAGATTTTGAAGCTGAAGCCCCGTATATTTAATTTATTTTGTAAACACAAAA

FIGURE 70

MTPSPLLLLLLPLLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTM
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDISP
ESLGPDSSSGGQEDPASQQWARPRFTQPSKMRRRVIAARPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRS DVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLT VLPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTL LLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTSHHTSHVEGKV
HQHIHYQC

FIGURE 71

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAACTTCCCAGGGGACCGCATTCCAGAGTC
 AGTGACTCTGTGAAGCACCCACATCTACCTCTTGCCACGTTCCACGGGCTTGGGGGAAAG**ATG**GTGGGGACCA
 AGGCCTGGGTGTTCTCCTTCTGGTCTTGAAGTCACTCTGTGTTGGGGAGACAGACGATGCTCACCCAGTCA
 GTAAGAAGAGTCCAGCCTGGGAAGAAGAACCCAGCATCTTTGCCAAGCCTGCCGACACCCTGGAGAGCCCTGG
 TGAGTGGACAACATGGTTCAACATCGACTACCCAGGCGGGAAGGGCGACTATGAGCGGCTGGACGCCATTTCGCT
 TCTACTATGGGGACCGTGTATGTGCCCGTCCCCTGCGGCTAGAGGGCTCGGACCACTGACTGGACACCTGCGGGC
 AGCACTGGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTTCTGGTGCCTCAACAGGGAGCAGCGGCCTGGCCA
 GAAGTCTCTAATTACACCGTACGCTTCCCTCTGCCACCAGGATCCCTGCGCCGAGACACAGAGCGCATCTGGA
 GCCCATGGTCTCCCTGGAGCAAGTGCTCAGCTGCCTGTGGTCACTGGGGTCCAGACTCGCACACGCATTTGC
 TTGGCAGAGATGGTGTGCTGTGCACTGAGGCGAGCGAAGAGGGTCACTGCACTGCATGGGCCAGGACTGTACAGC
 CTGTGACTGACCTGCCAATGGGCCAGGTGAATGCTGACTGTGATGCCTGCATGTGCCAGGACTTCATGCTTC
 ATGGGGCTGTCTCCCTTCCCGGAGGTGCCCCAGCCTCAGGGGCTGCTATCTACCTCCTGACCAAGACGCCGAAG
 CTGCTGACCCAGACAGACAGTGTATGGGAGATTCGGAATCCCTGGCTTGTGCCCTGATGGCAAAAGCATCCTGAA
 GATCACAAAGGTCAAGTTTGCCCCATTGTACTACAATGCCAAGACTAGCCTGAAGGCAGGCATCATGCAAG
 CAGAGTTTGTGAGGGCAGAGACTCCATACATGGTGTATGAACCTGAGACAAAAGCACGGAGAGCTGGGCAGAGC
 GTGTCTCTGTGCTGTAAAGGCCACAGGGAAGCCAGGCGAGACAAGTATTTTTGGTATCATAATGACACATTGCT
 GGATCCTTCCCTCTACACGATGAGAGCAAGCTGGTGTGAGGAACTGCAGCAGCACCAGGCTGGGGAGTACT
 TTTGCAAGGCCAGAGTGTGCTGGGGCTGTGAAGTCCAAGGTTGCCAGCTGATTGTACAGCATCTGATGAG
 ACTCCTTGAACCCAGTTCCTGAGAGCTATCTTATCCGGCTGCCCATGATTGCTTTCAGAATGCCACCAACTC
 CTTCTGACTATGACCTGGGACGCTGCCCTGTTAAGACTTGTGACGGGCAGCAGGATAATGGGATCAGGTGCCGTG
 ATGCTGTGCAGAACTGCTGTGGCATCTCCAAGACAGAGGAAAGGGAGATCCAGTGCAGTGGCTACACGCTACCC
 ACCAAGGTGGCCAGAGGATGCAGCTGCCAGCGGTGTACGGAACTCGGAGCATCGTGCGGGGCCGCTGTCACTGC
 TGTGACAATGGGGAGGCCATGCGCTTTGGCCATGTATCATGGGGAACAGCCGTGTAAGCATGACTGGCTACA
 AGGGCACTTTACCCCTCCATGTCCCCAGGACACTGAGAGGCTGGTGTCTACATTTGTGGACAGGCTGCAGAAG
 TTTGTCAACACCACCAAGTGTACCTTTCAACAAGAAGGGGAGTGGCGTGTCCATGAAATCAAGATGCTTCG
 TCGGAAAGAGCCCATCACTTTGGAAGCCATGGAGACCAACATCATCCCCCTGGGGGAAGTGGTTGGTGAAGACC
 CCATGGCTGAAGTGGAGATTCCATCCAGGAGTTTCTACAGGCAGAATGGGGAGCCCTACATAGGAAAAGTGAAG
 GCCAGTGTGACCTTCTTGGATCCCCGGAATATTTCCACAGCCACAGCTGCCAGACTGACCTGAACCTTCATCAA
 TGACGAAGGAGACACTTTCCCCCTTGGACGTATGGCATGTTCTCTGTGGACTTCAGAGATGAGGTCACTCAG
 AGCCACTTAATGCTGGCAAAGTGAAGGTCCACCTTGACTCGACCCAGGTCAAGATGCCAGAGCACATATCCACA
 GTGAAACTCTGGTCACTCAATCCAGACACAGGGCTGTGGGAGGAGGAAGGTGATTTCAAATTTGAAAATCAAAG
 GAGGAACAAAAGAGAGACAGAACCTTCTGTGGGCAACCTGGAGATTCTGTGAGAGGAGGCTCTTTAACCTGG
 ATGTTCTGAAAGCAGGCGGTGCTTTGTTAAGGTGAGGGCCTACCGGAGTGAGAGGTTCTTGCCTAGTGAGCAG
 ATCCAGGGGGTGTGATCTCCGTGATTAACCTGGAGCCTAGAATGGCTTCTTGTCCAACCTAGGGCCCTGGGG
 CCGCTTTGACAGTGTATCACAGGCCCAACGGGGCTGTGTGCTGCCTTCTGTGATGACCACTCCCTGATG
 CCTACTCTGCCTATGTCTTGGCAAGCCTGGCTGGGAGGAACTGCAAGCAGTGGAGTCTTCTCCTAAATTCAC
 CCAATGCAATTTGGCGTCCCTCAGCCCTATCTCAACAAGCTCAACTACCGTCCGACGGACCATGAGGATCCACG
 GGTAAAAAAGACAGCTTTCCAGATTAGCATGGCCAAAGGCCCAAGGCCCACTCAGCTGAGGAGAGCAATGGGCCCA
 TCTATGCCTTTGAGAACCTCCGGGCATGTGAAGAGGCACCCAGTGCAGCCCACTTCCGGTTCTACCAGATT
 GAGGGGGATCGATATGACTACACACAGTCCCTTCAACGAAGATGACCCTATGAGCTGGACTGAAGACTATCT
 GGCATGGTGGCCAAAGCCGATGGAATTGAGGGCTGCTATATCAAGGTGAAGATTGTGGGGCCACTGGAAGTGA
 ATGTGCGATCCCGCAACATGGGGGCACTCATCGGCGGACAGTGGGGAAGCTGTATGGAATCCGAGATGTGAGG
 AGCACTCGGGACAGGGACAGCCCAATGTCTCAGCTGCCTGTCTGGAGTTCAAGTGCAGTGGGATGCTCTATGA
 TCAGGACCGTGTGGACCGCACCTGGTGAAGGTCAATCCCCAGGGCAGCTGCCGTGAGCCAGTGTGAACCCCA
 TGCTGCATGAGTACCTGGTCAACCACTTGCCACTTGCACTCAACAACGACACCACTGAGTACACCATGCTGGCA
 CCGTTGGACCCACTGGGGCCACAATATGGCATCTACACTGTCACTGACCAGGACCCTCGCACGGCCAAAGGAGAT
 CGCGCTCGGCCGGTGTCTTGTGATGGCACATCCGATGGCTCCTCCAGAATCATGAAGAGCAATGTGGGAGTAGCCC
 TCACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGAGTGCCTTCCAGTACCTCCAAAGCACCCAGCCAG
 TCCCTGTCTGCAGGCACTGTCCAAGGAAGAGTGGCTCGAGGAGGCAGCAGCGAGCGAGCAGGGGTGGCCAGCG
 CCAGGGTGGAGTGGTGGCTCTCTGAGATTTCTAGAGTTGCTCAACAGCCCTGATCAACT**TAAG**TTTTGTGGT
 ACTTCACCTCTTCTGCCCTCATTTTCATGTGACAGCCATTGTGAGACTGATGCACAACTGTCACTTGGTTAAT
 TTAAGCACTTCTGTTTTCTGTGAATTTGCTTGTGTTTCTTTCATGCCTTTACTTACTTTGTCCCATGCTACTGA
 TTGGCACGTGGCCCCACAATGGCACATAAAGCCCTTTGTGAAACTGTTCTTTAAATGAAACACAAGAAAT
 GGCCACTGGTAAACTCTGCAGCTTCAACTGTACTTCATTTAATGCCATTAATGCAATATACTTCTCTTCTT
 TTTGCATGGTTTTTGGCCACCTCTGCAATAGTGATAATCTGATGCTGAAGATCAAATAACCAATATAAAGCATAT
 TTCTTGGCCTTGTCTCCACAGGACATAGGCAAGCCTTGATCATAGTTCATACATATAAATGGTGGTGAATAAAG
 AAATAAAACACAATACTTTTACTTGAATGTAAATAACTTATTTATTTCTTTGCTAAATTTGGAATTTCTAGTGC
 ACATTCAAAGTTAAGCTATTAAATATAGGGTGATCATAGTTCTCTACCAAGTCTGGAAAGAACATCTCCTGGT
 ATCCACAATTACACAGGTTGCTAAGTATTTGTACATTTCCCTTTGCATTTCGCTTTTGTCTTGCTAGAAAC
 CCAGTGTAGCCAGGGCAGATGTCAATAAATGCATACTCTGTATTTGAAAAA

FIGURE 72

MVTGKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSTFAKPADTLESPGEWTTWFNI
 DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ
 RPGQNCSNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS
 EASEEGQHCMGQDCTACDLTCMPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAYLLTR
 TPKLLTQTDS DGRFRI PGLCPDGKSILKITKVKFAPIVL TMPKTS LKAATIKAEFVRAETPY
 MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRLKQQHQAG
 EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLI RLP HDCFQ NATNSFY YDVGRCPV
 KTCAGQQDNGIRCRDAVQNC CGISKTEERE IQCSGYTLPTKVAKECSCQRCTETRSIVRGRV
 SAADNGEPMRFGHVYMGNSRVSM TGYKGTFTLHVPQDTERLVLTFVDRLQKFVN TTKVLPFN
 KKGS AVFHEIKMLRRKEPIT LEAMETNII PLGEVVGEDPMAELEIPSR SFYRQNGEPYIGKV
 KASVTFLDPRNISTATAAQ TDLNFINDEGDTFPLRTYGMFSVD FRDEVTSEPLNAGKVKVHL
 DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL
 DVPESRRCFVKVRAYR SERFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA
 CVP AFCD DQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPR
 VKKTA FQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN
 EDDPMSWTEDYLAWWP KPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRS
 TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTL VKVIPQGS CRRASVNPMLHEYLVNHLPLAV
 NNDTSEY TMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT
 FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVP SRRQQRASRGGQRQGGVVASLRFPRVA
 OOPLIN

FIGURE 73

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCAC
AATATCTTAACCTTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCATCTTCATTTAAAAAAAATACAGAG
ACCTACCTACCCGTACGCATACATACATATGTGTATATATATGTAAACTAGACAAAAGATCGCAGATCATAAAGC
AAGCTCTGCTTTAGTTTTCCAAGAAGATTACAAAGAATTTAGAGATGATTTGTCAAGATCCCTGTCGATTTCATG
CCCTTTGGGTTACGGTGTCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACT
CAGATTTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCTGCCAGCCGGAATCCACGGACATGACAAAATA
TCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC
CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCCTGAGCTGATGTTTGATTTT
GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT
CACTCTGTCTTGGAGCAAACCATTTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC
AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA
GATGCTTTTTCATATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTGTCACAGA
AGAGTACTCAACAGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTCCGCGCTTTTGT
CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAAACTCAGAGATTTCTTT
ACAGTCACAGACCTGAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC
ACGCTACTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAATCTCCATGCCACTGTATGTG
TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACTACAGGTCCAGACTGTGGGAAATGCAAGAAG
AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAATACCTGTATCCC
CAGTATTTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA
ACAACGTGCGCTGCCTGTGCCCCGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGTGGC
AGCTGCGGCTCCGACTCTGGCCAGGGCGCCCCGACGGCACCCCGAGCGCTGCTGCTGCTGACCACGCTGCT
GGGAACCGCCAGCCCCCTGGTGTTC**TAGGT**GTCACCTCCAGCCACACCGGACGGGCCTGTGCCGTGGGGAAGCA
GACACAACCCAAACATTTGCTACTAACATAGGAAACACACACATACAGACACCCCCACTCAGACAGTGTACAAA
CTAAGAAGGCCTAACTGAATAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTT
TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA
TTGGAAAGGCTGCGACAGCCCCCAAACAGGAAAGACAAAAACAAACAAATCAACCGACCTAAAAACATTTGGC
TACTCTAGCGTGGTGCGCCCTAGTACGACTCCGCCCAGTGTGTGGACCAACCAATAGCATTCCTTTGCTGTGAG
GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC
CTGTGCTTTAGTGAACGTTGCTCTGTAACCCTCGTTGGTTGAAAGATTTCTTTGTCTGATGTTAGTGATGCACA
TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCAG
GCACACACCCACTATACAAGAGTGGCTATAGGAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTT
ATTTTTCTTGAACACTGTAATATGTAGATTTTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAAT
GTATCTAATTCGAATCAGCAAAGACTGACATTTTATTTTGTCTCTTTCGTTCTGTTTTGTTTCACTGTGCAGA
GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATATAACAAGTGAATAAGA
TTCCACCAAAGGACATTCTAAATGTTTTCTTGTGCTTTAACACTGGAAGATTTAAAGAATAAAACTCCTGCA
TAAACGATTTCAGGAATTTGTATTGCAATTTCTTAAGATGAAAGGAACAGCCACCAAGCAGTTTCACACTCACT
TTACTGATTTCTGTGTGGACTGAGTACATTCACTGACGAATTTAGTTCCAGGAAGATGGATTGATGTTCACT
AGCTTGGACAACCTCTGCAAAATATGAGACTATTTCACTTGGGAAAAATTACAACAGCAAAAAAAAAAAAAA
AAAAAA

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK
VKLDPPDITCGDPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMLEKSLDYGRTWQPYQYYATDCLDAF
HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKI IHFEIKDRFALFAGPRLRNMASLYGQLD
TTKKLRDFFTVTDLRIRLLRPVAGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPI PKGTANTCIPSISSIGTNVCDNELLH
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS
PLVF

FIGURE 75

CCACGCGCTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCCGGCTAAGGATTGCTGAGGAGGCGC
CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCCGTCCGGGCGAGGTGTCTCATGACTT
CTCTTGTGGACCCATGTCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC
CCTCTCAGCCTCTACTGATTTTTTACCACACCCAAGATTTTTTGGAAATGGAGGAGACGGCTCA
AGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTT
AGTATACATTTTTCTTCTTTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC
AGCAGCCATGGCCTTCTGCTTCTTGAGACCCTGTGGTGGGAATTCACAGCTTCTATGACA
CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTTCTTGAGTTTGACAGCATCATTAG
AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAATAAAT
TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA
ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA
CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT
CATTTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT
TGGACCAAACCTCGTTGAGCCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT
CCGGGAGCAGTGATGTCAAACCTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA
AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTTCGGAAACATCTGCCCATGTGTATTG
ATGGCAGAGCTGTTGCCACAAAGCGCCTTTTATTTAGGGTAAATTAACAAATCCATTCTAT
TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT
GCTTCAGAAGTGTTATTTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT
TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAACCAT
TTACTTTACATATTCGTTTTCAATACTTGCTGTTTCATGTTACACAAGCTTCTTACGGTTTTT
TTGTAACAATAAATATTTTGGAGTAAATAATGGGTACATTTTAAACAAACTCAGTAGTACAACC
TAAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAAATG
AGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTTTATTAAAAAAAAAAAAAAAAAAAAA
AAAAG

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFS IHF
SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKV KW
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

Downloaded from www.ashg.org

FIGURE 77

TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTGAGAAAGTGAAGTGGCATT
TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGGAAAAAATTCAGGAGGAGCTCAAG
TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT

030304.1.1301

FIGURE 78

CTCAGCGGCGCTTCTCCTCGTAGCGAGCCTAGTGCGGGGTGTTTGCAATTGAAACGTGAGCGCGA
CCCGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAACGAGGCGGGTGGT
CCTGCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCGAGTT
TCTGTGCGAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGGTGCTTGGCGGCGGCGGCTT
CCTCCCCGCTCGTCTCCCCGGGCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTA
TGGAAGCACCTGACTACGAAGTGCTATCCGTGCGGAAACAGCTATTCCACGAGAGGATCCGC
GAGTGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC
CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA
TTGCGCTCGAGCTGTGCACCTTTACCTGGCAATTGCCCTGGGTGCTGTCTCTGCTCCTGCC
TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAACTACTACATCCAGTGGCT
CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCAACCTGTCCCTCA
TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG
GGTGTCTTGGGCCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT
AGGTATGGTGTTGGGTGGCATCAGCCATTGTGGACAAGAAACAAGGCCAACAGAGAGTCACTCT
ATGACTTTTTGGGAGTACTATCTCCCCTACCTCTACTCATGCATCTCCTTCCTTGGGGTTCTG
CTGCTCCTGGTGTTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT
AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG
CAGCCCTGACCCGCAGGATCTGTAATCCTACTTCCTGCTGGCTGCCTTTAGACATGGAGCTG
CTACACAGACAGGTCTGGCTCTGCAGACACAGAGGGTCTGCTGGAGAAGAGGCGGAAGGC
TTCAGCCTGGCAACGGAACCTGGGCTACCCCTGGCTATGCTGTGCTTGCTGGTGCTGACGG
GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG
CCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG
TGCCGTCAATCAGGTTGTACTCATCTTTTACCTAATGGTGTCCTCAGTTGTGGGCTTCTATA
GCTCTCCACTCTTCCGGAGCCTGCGGCCAGATGGCACGACACTGCCATGACGCAGATAATT
GGGAACTGTGTCTGTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCCTGGG
GCTCACTCGCTTTGACCTGCTGGGTGACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACA
TTGTGTTCTCTACAACGCAGCCTTTGCAGGCCTCACCACACTCTGTCTGGTGAAGACCTTC
ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT
CTCCGGTTTTCCCCAGGCATCTAGGAAGACCCAGCACCAG**TGA**CTCCAGCTGGGGGTGGGA
AGGAAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG
ACCTCAGGACCTGGAATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATT
GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAAGTGTGGCCT
CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA
ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT
AGGGTGGGGTAGGGGCAGGGAAAGGACTGGGCCAGGGCAGGCTCGGGAGATAGATTGTCTCC
CTTGCCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG
AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCAGGGA
AAAAAA

FIGURE 79

MEAPDYEVL SVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATV NK
IALELCTFTLAIALGAVLLL PFSIISNEVLLSLPRNYI IQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGV LGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRILLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRV LLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMP RGMQGTSLGQVSFSKLGSFGAVIQVVLIFYLMVSSVVG FY
SSPLFRSLRPRWHD TAMTQIIGNCVCLLVLSALPVFSRTLGLTRFDLLGDFGRFNWLG NFY
IVFLYNAAFAGLTTLCLVKTFTA AVRAELIRAFGLDRLPLPVSGFPQASRKTQH Q

FIGURE 80

GGCTGCCGAGGGAAGGCCCTTGGGTTGGTCTTGGTTGCTTGGCGGCGGCGGNTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

FIGURE 81

GACCGACCTTAAAGAGTGGGAGCAAAGGGAGGACAGAGCCTTTTAAACGAGGCGGTGGTGC
CTGCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTC
TGTCGCAGGCTGCGAGGAAAGGCCCCCTAGGCTGGGTCTGGTGCTTGGCGGCGGCGGCTTCCT
CCCCGTTGTCNTCCCCGGGCCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGA
AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGT
GTATTATATCAACACTTCTGTTTGCAACACTGTACATCNTCTGCCACATCTTCCTGACCCGC
TTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGATTGC
GCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCTCCTGCCCTTCT
CCATCATCAGCAATGAGGTGCTGCACTCCC

FIGURE 82

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTT
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTT**CATG**CTGCTGT
GGGTGATATTACTGGTCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCAT
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA
GGGATTTTCGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGAAA
TACTAAGAGAAACCCAGACAATATCCTTGAGGTTCAAGGAATCTGGAGAGTACAGATGCCAG
GCCCAGGGCTCCCCCTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCAC**CTAG**GCCTCTC
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA
TGATAATGTCCTGGCATTCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 83

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSQKTKWYHRYL
GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT

FIGURE 84

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGTGT
GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG
GAGTGTGTGGAACAGGACCCGGGACAGAGGAACC**ATG**GCTCCGCAGAACCTGAGCACCTTTT
GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG
GGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAACTAGCCCTGCA
GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCAGGATCTGGGTG
CTGCTTATGAGGTTCTGTCTAGATAGTGAGAAACGGAACAGTACGATACTTATGGTGAAGAA
GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCACACTTCTTTGGGGATTT
TGGTTTCATGTTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATA
TTATTGTAGATCTAGAAGTCACTTTTGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT
AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTGCGCAAGAGAT
GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT
GCCCTAATGTCAAACCTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG
AGAGACGGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG
AGATTTACGGTTCCGAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATT
TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT
CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT
ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA
TCACTTTTGTATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA
CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TGA**GAGTG
AATAAAATTGGACTTTGTTTTAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTT
TTGTGTGTGTTTTTTGTTTTTATTTTCAATATGCAAGTTAGGCTTAATTTTTTTTATCTAATGA
TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTTCGGAAAAGAATGACC
AGCAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT
TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA
GTTGTTAGCAATTTTCAATCAAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTG
TTATTTTTTA

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
 AQEKFQDLGAAYEVLSDSEKRKQYDTYGEGLKDGHQSSHGDI FSHFFGDFGFMFGGTTPRQQ
 DRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
 MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGM EYPFIGEGEPHVDGEPGDLRFRIKVVKH
 PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
 NNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

FIGURE 87

GGCACGAGGCGGCGGGGAGTTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATGG**
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC
TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC
TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC
AGAGAAGCTTGTGGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG
TACCCTCCGTTGGACCCCCAACTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTCAG
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC
AGTCTCTGTGCGCTGCTGAGGAGCATTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATT**TTAGT**GCCT
ACAGGCCAGCAGCTAGCCATGAAGGCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG
TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAGTTGCAAACCTGTGGCTGGTGAGT
GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT
AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG
TGTGGCAATTCTGATCTGCATTTTCAGAAGAGGACAATCAATTGAACTAAGTAGGGGTTTC
TTCTTTTGGCAAGACTTGTAATCTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAAGTTTGGGTTTGAAGCTGAGGAACT
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG
GATAGTAAATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAA

FIGURE 88

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL
ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS
VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLSVSHLVLVTRNACHLTGGLDWI
DQSLSAAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

TOPT-140650

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTCCGAGGTGCTTTCGCCGCTGTCC
 CCACCACTGCAGCC**ATG**ATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA
 TTTGGAGTGTTTTTCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT
 TGGAAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAATTGGTTTAGAAAGAACATTCAGAT
 TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC
 CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTTCAG
 GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT
 TTACCTGGAATTAGATCATTGTAGATAAAGTTGGAGAAAGCAACAATATGGTA**TAA**CAACA
 AGTGAATTTGAAGACTCATTTAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA
 GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAACGTATAG
 CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA
 ACTAAGAAGAAGTCAGCAAGCAAAGTGAAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA
 ACTCTTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA
 CTGTGGTGCCTGTTTCTTTCTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT
 TTTTAGAAGTGCTCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA
 TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG
 ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

FIGURE 90

MISLTDTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK
HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFI RRVPVLGSLLNLPGI
RSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

FIGURE 91

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTTCCTGTTCTTTGGA
ATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT
GGCTTTTGTAATTGGTTTAGAAAGAACATTTCAGATTCTTCTTCCAAAAACATAAAATGAAAG
CTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTT

FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA
GGCTGCCAGGAAGGAGACGCCTTCCTGAGTCCTGGATCTTTCTTCCTTCTGGAAATCTTTGA
CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**ATG**GACCTCGCGGGACTGC
TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC
ATCAACACCATTTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA
GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT
CGGGCACGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAAT
GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA
ACGCTTTGGGCTGTTAGGGGGCTCCAAGGTCTGGCCAAGAAAGAGCTGGCCTATGTCCCAA
TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTTCGCGCAAGTGGGAGCAGGAT
CGAAGACGGTTGCCACCAGTTTGAGCAGCCTCCGGGACTACCCCGAGAAGTATTTTTTTCCT
GATTCACTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC
GGGCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC
ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTCAGAAA
TAATGAAAATCCAACACTGCTGGGAGTCTTAAACGGAAAGAAATACCATGCAGATTTGTATG
TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC
AAGCTCTACCAGGAGAAGGATGCCTTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCGAGA
GACGCCCATGGTGCCCCCCCCGGCGGCCCTGGACCCTCGTGAACCTGGCTGTTTTTGGGCCTCGC
TGGTGCTCTACCTTTTCTTCCAGTTTCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG
CTGGCCAGCTTCATCCTCGTCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGT
GACGGAAATTGACAAGGGCTCTGCCTACGGCAACTCTGACAGCAAGCAGAACTGAATGAC**T**
GACTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAACTGGTGGCCTCTGCATATCCT
CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA
CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT
TTTTCCCCATGTGCTTTAGTGGGCTTTGGTTTTCTTTTTTGTGCGAGTGTGTGTGAGAATGGC
TGTGTGGTGAGTGTGAACCTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG
GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT
AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC
CAAGAAAAAAAATTAAGTGCTTTTCTGGGTCAAAAAAAAAAAAA

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIA SGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV
MLLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHNKFEIDFLCGWSLSERFGLLGSKVLAKK
ELAYVPIIGWMWYFTEMVFC SRKWEQDRKTVATSLQHLDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKK
YHADLYVRRIPLEDIPEDDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRRMIGVTEIDKGSAYGNSDS
KQKLND

"T406666"

FIGURE 94

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG
 GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA
 GTAAAAGGTGAAGCCAAGAAGCAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA
 TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTATAGCTTTTATAATTCTTCAG
 GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT
 TGGTACAAATTCCGTCGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA
 AAAGTTGCAGGAGCATTTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA
 TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT
 TTTTACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAAGTGGTTATAAAAC
 TGTATCAGGTTCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT
 TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA
 CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT
 AGTAAAGGATGTAAACAGATTAACACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG
 CAGCAAGAGAGAAGAACATCCAAAAAGACCTCAGGAGAACATTTTTCTTTGTCAGGCATTA
 CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTTCATGTGTTATGTCTTTAAAAAATAGACA
 TGTTTCTAAAAGTAGCTGTAACACCAACCATCTCGATGTAGTAGACAATCTGACCTTAA
 TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT
 AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTGTTAGATACACAAGA
 CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC
 CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA
 TTT**TGA**TCCTTTTAAACCTTACAAGGAGATTTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT
 TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTTCAATTTGTTTTTACTATGTTTAC
 CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTTCAAAC
 ATCAGATGCTTTTTATTTCCAAACCTTTTTTTTACCTTTCACCTTCACTAAGTTGTTGAGGGGAAGGCT
 TACACAGACACATTCTTTAGAAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA
 TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC
 TGGGCAACGTATTGAGACCATGTCTATTAATAAATAAATGGAAAAGCAAGAATAGCCTTAT
 TTTCAAATATGGAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG
 TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAATGCTGGATATCATGCA
 ATAAATTTGCAAAACATCATCTAAAATTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLLDTDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

FIGURE 96

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGC
 CCAAGCAGCGCGCAGCGAACGCCCCGCGCCGCCCCACACCCTCTGCGGTCCCCGCGGCGCTGCCACCCTTCCCT
 CCTTCCCCGCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCCCGAGGTACCA
 GCCCGCGCTCTGCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCCCTGGCCCCGGCGCTGGCACC
 GGGGACCGTTGCTGACGCGAGGCCCCAGCTCTACTTTTTCGCCCCGCTCTCCTCCGCTGCTCGCCTCTTCCAC
 CAACTCCAACTCCTTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCCGCTGCCGTAG
 CGCCGCTTCCCGTCCCGTCCCAAAGGTGGGAACGCGTCCGCCCGGCCCGCACCATGGCACGGTTCGGCTTGCC
 CGCGCTTCTCTGCACCCTGGCAGTGCTCAGCGCCGCGCTGCTGGGTGCCGAGCTCAAGTCGAAAAGTTGCTCGG
 AAGTGCAGCTCTTTACGTGTCCAAAGGCTTCAACAAGAAGCATGCCCCCTCCACGAGATCAACGGTGATCAT
 TTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA
 TGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTGCAAGCTGTCTTTGCTTACGTTACAAGAAGTTTG
 ATGAATTCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT
 TTATACATGCAAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT
 GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCTGGTGAACCTCCAGT
 ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT
 CGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCCGTACTTTGCTCAAGGCTTAGCGGTTGCGGG
 AGATGTGCTGAGCAAGGTCTCCGTGGTAAACCCACAGCCCAGTGTACCCATGCCCTGTTGAAGATGATCTACT
 GCTCCCACTGCCGGGCTCTCGTGACTGTGAAGCCATGTTACAACACTACTGCTCAAACATCATGAGAGGCTGTTTG
 GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA
 GGGTCCTTTCAACATTGAATCGGTTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG
 ATAATAGTGTTCAAGTGCTCAGAAGGTTTCCAGGGATGTGGACCCCCAAGCCCCCTCCAGCTGGACGAATT
 TCTCGTTCCATCTCTGAAAGTGCCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCCAACACAGC
 AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTCTGGTCCT
 CCCTTCCGAGCAACGTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG
 AAAGGCCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCA
 GGTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA
 AGAATGCATACAATGGGAACGACGTGGACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAAAGT
 GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAAATGA
 GAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGG
 TTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTCTGAGAAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAGTT
 ATCACTTTTCTACCATCCTAGTGACTTTGCTTTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGC
 CACTGGTTTAAGAAGTGCTGACTTTGTTTTCTCATTAGTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGT
 TCCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAACTATAGTTAGTTGTGCATTTGTGA
 TTTTATCACTCTATTATTTGTTTGTATGTTTTTTCTCATTTCGTTTGTGGGTTTTTTTTTCCAAGTGTGATCT
 CGCCTTGTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATA
 GCTGTACAGAAGCAGGTTTTATTATCATGTTATCTTATTAAAAGAAAAAGCCCCAAAAGC

FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAAEKSLNDMF
VKTYGHLYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFLVNSQYHFTDEY
LECVSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGC GPPKPLPAGRISRSISESAFSARFRPHHPEE
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLE
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

FIGURE 98

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA
GCAACTTACAGCTGCACCGACAGTTGCG**ATG**AAAGTTCTAATCTCTTCCCTCCTCCTGTTGC
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCAC
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA
AGATTGGTTCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGC
AGTGCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGA
AAGCCAAACAAGCATTCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT
TGCTCTGCCTTTG**TAG**GAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCACCTCACTCTCCCACTGTACCCACC
CCTAAATCATTCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTC
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT
AATTACCTGAAAGATTCAGGAAACTGTAGCTTCCTAGCTAGTGTCAATTAACCTTAAATGC
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTTAAATGTCAAAAAAAAAAAAAAAAAA

99/330

FIGURE 99

MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR
KFMTVSGLPKKQCPDHFKNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Figure 1 consists of 15 small plots arranged in a grid, each showing the relationship between a specific variable and the probability of a 'Yes' response. The variables on the x-axis are: Age, Gender, Education, Income, Marital Status, Religion, Ethnicity, Occupation, Health Status, Lifestyle, Social Network, Cultural Values, Political Views, Economic Outlook, and Environmental Concerns. The y-axis for all plots is 'Probability of Yes', ranging from 0.0 to 1.0. The plots show various trends, such as increasing or decreasing probabilities with age, gender differences, and the impact of different factors on the likelihood of a 'Yes' response.

FIGURE 100

AATGGCTGTCTTAGTACTTCGCCTGACAGTTGTCCTGGGACTGCTTGTCTTATTCCTGACCT
GCTATGCAGACGACAAACCAGACAAGCCAGACGACAAGCCAGACGACTCGGGCAAAGACCCA
AAGCCAGACTTCCCCAAATTCCTAAGCCTCCTGGGCACAGAGATCATTGAGAATGCAGTCGA
GTTTCATCCTCCGCTCCATGTCCAGGAGCACAGGATTTATGGAATTTGATGATAATGAAGGAA
AACATTCATCAAAGTGACATCCTCAGGACACACCCATGTGGCTCCTGGACAATCCAAGAGCA
GCCAAATCCTGCTTTTCCAGTTTGGCTCCACAAGTCCTCCAGGACAGAGCCCTCAAAGCAAC
TCCCAACGAGTTCTCAGGATTCAGGCTCTGGCTTCAACCAAACAGAACTCATTTTGAACACC
CTGACTGCATTTTGTCTTTAGAAAGTTAGAATAAATATGGCGCTTTGGGATCACATAGTTG
ATGGAGAGGAAA

101/330

FIGURE 101

MAVLVLRLTVVVLGLLVLF LTCYADDKPKDPDDKPD DSGKDPKPDFPKFLSLLGTEIIENAVE
FILRSMSRSTGFM EFDDNEGKHSSK

099304.1.1.1.1

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT
 CAGAGCTGGTCTGCC**ATG**GACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTAC
 CCTGCCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAAGCTACTTCC
 CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG
 CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG
 CTGCGGAACCGGAGCCAACCTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC
 CAAATCCCCACTTTGAGAAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT
 GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT
 GGTGGTCTGCACTCTGGTGCTGTGCTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCC
 GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA
 AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG
 CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG
 AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC
 AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA
 AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACC**TAG**CAGAATGAGAGAAGACATT
 CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC
 CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC
 TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC
 CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT
 CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT
 CTCCCCACTACCACCTTCTTCCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT
 GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA
 AACCACG

FIGURE 103

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFFPYLMAVLTPKSNRKMESKKRELFSSQI
KGLTGASGKVALLELGC GTGANFQFYPPGCRVTCLDPNPHFEKFLT KSMAENRHLQYERFVV
APGEDMRQLADGSMDVVVCTLVLC SVQSPRKVLQEVRRVLRPGGVLFFWEHVAEPYGSWAFM
WQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWL PVGPHIMGKAVKQSFP
SSKALICSFPSLQLEQATHQPIYLP LRG T

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGGAAGTTGCCTCATCGCAGG
 CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA
 ACTAATATTTATATGACAGAAGAAAAAGATGTCATTCCGTAAAGTAAACATCATCATCTTG
 TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG
 TTAAGGAATGAGGTTACAGATTGAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAA
 TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCCTGTGGTCATCGCTGCATCTG
 AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTGAGCACAACACTCGCTCCAAT
 GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCTGGCTCAACAG
 TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACCCTAACTTTTGGAAGGAA
 AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC
 TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA
 AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG
 AAGATTGTGATTGAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC
 ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCATGAAAGCCAGCACTTG
 CTCATTTAATCCTGGAGTTTTTGTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA
 ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCTGGCT
 GGTCAGCATCACAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC
 TATGTGGAATGTCCGCCACCTTGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTA
 AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT
 ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTCACCTAATCCGAAG
 ATATACCGAGATCTCAAACATAAAGTGAACAGAATTTGAACTGTAAGCAAGCATTTCTCAG
 GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA
 GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTC
 AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTCCTT
 ACTACAATGCTGAATGACTGGAAAGAAGAACTGATATGGCTAGTTCAGCTAGCTGGTACAGA
 TAATTCAAAACTGCTGTTGGTTTTAATTTTGTAACCTGTGGCCTGATCTGTAAATAAACTT
 ACATTTTTC

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR
QEEIPVVIAASEDRLGGAIAAINS IQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDI LALYNT
ALKPGHAAAFSEDCDSASTKV VIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA
NLTEWKRQNITNQLEKWMKLNVEEGLYSRTLGSITTPPLLIVFYQQHSTIDPMWNVRHLGS
SAGKRYSPQFVKA AKLLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

FIGURE 106

TGGTTTTTGCCCCATAAATTCCCTCAGCTTGAGCAGTTTGTTAAGGAATGAGGTTACAGATT
CAGGAATTNTAGGNCCTCAACCTNTAGANTTTGTCCCAAATGTTCTCCGACATGCAGTAGAT
GGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATNTGAAGACAGGCTTGGGGGGGCCAT
TGCAGCTATAAACAGCATTTCAGCACAACTCGNTCCAATGTGATTTTCTACATTGTTACTC
TCAACAATACAGCAGACCATNTCCGGTCCTGGNTCAACAGTGATTCCCTGAAAAGCATCAGA
TACAAAATTGTCAATTTTGACCCTAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCA
GGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCG
CAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTAC
AATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTAC
TAAAGTTGTCATCCGTGGAGCAGGAAA

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG
GAGCTAGAGGGCAAGTGTGCTCGGCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC
TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTCCGGGGCGGGGCTGCA
TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT
TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT
TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT
CCGTAAAGTAAACATCATCATCTTGGTCCTGGGCTGTTGCTCTCTTCTTACTGGTTTTTGCAC
CATAACTTCCTCAGCTTGAGGCAGTTTGTTAAGGAATGAGGTTACAGATTCAGGAATTGTAG
GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA
GGAGATTCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA
ACAGCATTCAGCACAACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA
GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG
TCAATTTTGACCCTAACTTTTGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC
ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCACGCGCAAAGAAGG
CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA
CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTTCAGCCTCTACTAAAGTTGT
CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA
TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC
CTGACGGAATGGAACGACAGAATATACTAACCAACTGGAAAAATGGATGAACTCAATGT
AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAACACCTCCTCTGCTTATCG
TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT
GCTGGAAAACGATATTCACCTCAGTTTGTAAGGCTGCCAAGTTACTCCATTGGAATGGACA
TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA
GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAACATAAAGTGAAA
CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG
TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG
GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAGACTATAGACTATAAAT
ATGTCTCCATCTGCCTTACCAAGTGTCTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA
CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAACTGCTGTTGGTTTTAATTTT
GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTTCAATAGGTAAAAA

FIGURE 108

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG
TCTGTGCTGGTCTGAGGGTGTCTGCCTGT**CATG**GGGGGCAGCCATCTCCCAGGGGGCCCTCATC
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTCATCCTCTGCTG
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC
TGGCCCCCTGTCCTGAGAAGGCCCCACCACCCCAGAAGCCCAGCCATGAAGGCAGCTACCTGC
TGCAGCCCTGAAGGCCCCCTGGCCTAGCCTGGAGCCCAGGACC**TAA**GTCCACCTCACCTAGAG
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT
TGGAGCTGGACCCAGCGGCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCCGGGTCCACTCTTCCCTAGGCTGAGCACC
TCTAGGCCCTCTAGGTTGGGGAAGCAAACCTGGAACCCATGGCAATAATAGGAGGGTGTCCAG
GCTGGGCCCCCTCCCCTGGTCCTCCCAGTGTTTGCTGGATAATAAATGGAACCTATGGCTCTAA
AAAAAAAAAAAAAAAAAAAA

109/330

FIGURE 109

MGAAISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRRPHH
PRSPAMKAATCCSPEGPWPSLEPRT

109/330

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYDTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

PDB: 1D56

[illegible]

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQ
RQDPGAAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNSRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTLOPLVKRVCDTDRACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPD
LNAIKRVGYWSGGTSTGAAINFALQFLFKSKPNKRKLMILITDGRSYDDVRI PAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRI IQNICTEFNSQPRN

FIGURE 114

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCTGCGGGGCGCTGAGAGGACACGAGCTCTA
TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT
 GCGCCAGGTCCCGACGGCTCCGCGCCAGATCCCGCCCACTACAGTTTTTCTCTGACTCTAAT
 TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATTCCAAAGAGTGGTTGAAG
 TGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTTGAACAAACATT
 CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA
 GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACCTCC
 TCCCAGCCTTTCAGACCCCCACTGGCATGCCATATGGAACAGTGAACTTACTTCATGGCGTG
 AACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC
 CACCCTGAGCAGCCTCACTGGTGACCCGGTGTTTGAAGATGTGGCCAGAGTGGCTTTGATGC
 GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCACATTGATGTGCTCACTGGC
 AAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT
 GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCTAGAGTATAACAAAG
 CCATCCGGAACCTACACCCGCTTCGATGACTGGTACCTGTGGGTTTCAGATGTACAAGGGGACT
 GTGTCCATGCCAGTCTTCCAGTCCTTGGAGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG
 AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG
 GGCTCCCGGAATTCTACAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGGCTACCCA
 CTTCCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCT
 CCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT
 TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG
 GCCGAGACTGTGAAATACCTCTACCTCCTGTTTGACCCAACCAACTTCATCCACAACAATGG
 GTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTGCATCCTGGGGGCTGGGGGGTACA
 TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG
 GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCTCAAACGGAGCAGGTC
 GAAATTTTCAGAAAAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC
 TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCA
 CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCT
 AGACTCCTCATTAACCACTGGATAATTTTTTTATTTTTATTTTTTTGAGGCTAAACTATAATA
 AATTGCTTTTGGCTATCATAAAA

FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLLSKKAGVEVEAGWPCSGPLLRMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLWVQMYKGTVSMPVFQSLEAYWPGLQSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNIPOGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

FIGURE 116

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
 GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
 ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
 GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
 ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC
 TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
 AACAGTGTAATCTGTGCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
 GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCAGTGTGACATC
 ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
 GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
 TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
 CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
 GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
 CATTCTGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
 GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
 GGTCTGTGCCACTGTTCTGTGAAAATGGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
 TGGTCTCCAGACACCTTGAAAATAACCAATTCACCCAGAGTTAATCAGCTGCAGAAGG
 GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
 CTCA**TAG**GTTTTCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
 ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
 GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
 TGAAGTGGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
 CTGGGAAAAGTGAATTCATCCCTTCGGTCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
 TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
 TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTGAG
 TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
 GGCTTGGAGAGCCACTTTCCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
 TGTGAGTTCACCTCAAGCCCAATGCCGGTGAGAGGGGAATGGCTTAGCGAGCTCTACAGT
 AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
 CATGAACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
 GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
 TGTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
 AAAAAAAAAA

FIGURE 117

MQTFMTVLEEIWTSLEFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHELLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS AW
SILKHPFN RNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQT FVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFM LILV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

FIGURE 118

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCCAC
ATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 119

CGGACGCGTGGGCCGCCACCTCCGGAACAAGCC**ATG**GTGGCGGCGACGGTGGCAGCGGCGTG
GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG
TCAACATCCGGGGCAAACCTGGTGTGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG
AATGTGGCCAGCGAGTGCGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG
AGACCTGGGCCCCCACCCTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG
AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCGCCGCACCTACAGTGTCTCATTTCCC
ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA
GACTTCTGGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG
TGGTAGGGGCTTGGGACCCAACCTGTGTGAGTGGAGGAGGTCAGACCCAGATCACAGCGCTC
GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTA**TAA**CCACCGCGTCTCCTCCTCCACCA
CCTCATCCCGCCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAG
AGACCCACTGACTCTCCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA
AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG
AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA
TAGAAGTATATCAAGCAATAATCTCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC
CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC
CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT
AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT
GTTATTTCTCCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA
AACAATACCTCACGATATAAAATAAAAAATGAAAGTATCCTCCTCAAAAA

120/330

FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVNVASECGFTDQ
HYRALQQQLQRDLGPHHFNVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG
AHPAFKYLAQTSKKEPTWNFWKYLVA PDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

120/330

FIGURE 121

CGGACGCGTGGGCGGGCCGGGACGCAGGGCAAAGCGAGCC**ATG**GCTGTCTACGTCGGGATGC
 TGC GCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGCTGGGGGCCCCGGGCCGCCCTCTCT
 CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG
 CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTCAAGGGTGCACCAAAAAGCATCTTA
 ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCC
 TTGGTCGTCCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA
 AGCTGCTTCTGGCCTCCTGAGCATTTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC
 CTAACCTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCCAGGCGGGCATCATTCTGGTG
 TCTGTGAACCCAGCCTACCAGGCTATGGAAGTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA
 GGCCCTTGTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT
 GTCCAGAAGTGGAGAATGCCAGCCAGGGGGCCTTGAAGAGTCAGAGGCTCCAGATCTGACC
 ACAGTCATCTCGGTGGATGCCCCCTTTGCCGGGGACCTGCTCCTGGATGAAGTGGTGGCGGC
 TGGCAGCACACGGCAGCATCTGGACCAGCTCCAATAACAACAGCAGTTCTGTCTGCCATG
 ACCCATCAACATCCAGTTACCTCGGGGACAACAGGCAGCCCCAAGGGGGGCCACCTCTCC
 CACTACAACATTGTCAACAACCTCAACATTTTAGGAGAGCGCCTGAACTGCATGAGAAGAC
 ACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCTGTACCATTGCCTGGGTTCCGTGGCAG
 GCACAATGATGTGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC
 AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT
 GTTCGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG
 GTGTCAATTGCTGGGTCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT
 ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCCTGACATTCGCGCACTT
 CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGG
 CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGC
 ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT
 GGATCAGGACAAGTGGTATTGGACAGGAGATGTCGCCACAATGAATGAGCAGGGCTTCTGCA
 AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG
 CTCGAGGACTTCTTTTACACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA
 CGATCGGATGGGGGAAGAGATTTGTGCCTGCATTTCGGCTGAAGGACGGGGAGGAGACCACGG
 TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC
 GTGTTTGTACAAACTACCCCTCACCATTTCAGGAAAGATCCAGAAATTCAAACCTTCGAGA
 GCAGATGGAACGACATCTAAATCTG**TGA**ATAAAGCAGCAGGCCTGTCCTGGCCGGTTGGCTT
 GACTCTCTCCTGTCAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC
 TGAGCCAGGCACATCAAATGTCAAGGAATTGACTGAACGAACATAAGAGCTCCTGGATGGGTC
 CGGGAACCTCGCCTGGGCACAAGGTGCCAAAAGGCAGGCAGCCTGCCAGGCCCTCCCTCCTG
 TCCATCCCCCACATTCCCCTGTCTGTCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTT
 GAAAAAAAAAAAAAAAAA

FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRWQEARLQGVRFLLSSREVDRMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLP GTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCCKIVGRSKDMIIRGGENIYPAELEDFHHTHPKVQEVQVVGVKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

Signal Peptide:

amino acids 1-22

Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

Putative AMP-binding Domain Signature:

amino acids 260-271

N-myristoylation Sites:

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277,
314-319, 318-323, 379-384, 380-385, 409-413

N-glycosylation Site:

amino acids 282-285

FIGURE 123

CAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGACACCAGAGCAGTTGCGGA
TGATCCTGCCCCAACCCCTGTACCATTCCTGGGTTCGGTGGCAGGCACAATGATGTGTCTG
ATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGCAAGAAGGCACTGGAGGC
CATCAGCAGAGAGAGAGGACCTTCCTGTATGGTACCCCCACGATGTTTCGTGGACATTCTGA
ACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAGGTGTCATTGCTGGGTCC
CCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAATATGAAGGACCTGGTGGT
TGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTTCCCTGAGGACACTGTGG
AGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGGCGCGGATCATGAACATG
GAGGCAGGGACGCTGGCAAAGCTGAACACGCCCGGGGAGCTGTGCATCCGAGGGTACTGCGT
CATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGTGGATCAGGACAAGTGGT
ATTGGACAGGAGATGTCGCCAC

FIGURE 124

GAGCAGGACGGAGCC**ATG**GACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGC
 AGGCTGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG
 TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC
 GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC
 AGTGCGGGGTTGCGGTTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC
 TTCTGGCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC
 ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCAACGGCGTGGAGTG
 CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTTCGTGAGCT
 GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTACCTTGACGGCA
 GCTAATGTGACTGTGTCCTTGCCCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA
 TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAAC
 CTGACCTCCGCAACAAGACCTACTTCTCCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT
 CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCAACCACTTCTACCTCGGCCCCAGTGAG
 ACCCACATCCACCACCAAACCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG
 AACACGAGGCCTCCCGGGATGAGGAGCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC
 CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG
 TGTGGCTCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTACTGT**T**
GAGCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT
 CATCACTTCCTGTTCCCACTACTGGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCC
 CAGTATCCCCAGCTTCTGCTGCGCTGGTTTTCGGCTTTGGGAAATAAAATACCGTTGTATAT
 ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC
 TCCGCTTGTCTCTTGTGATGTTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTG
 AGAGAGAGGATGCTAAGCTTCCTACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGG
 GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG
 GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC
 CCAATTCGCCCTATAGTGAGTCGTA

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCS PNKMKTVKCAPGVDVCT
EAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL
DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTV
SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRI PPLVRLPPPEPTT
VASTTSVTTSTSAVRPTSTTKMPAPTSQTTPRQGEHEASRDEEPRLTGGAAGHQDRSNSG
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

099044-1-01
10511-1-01

FIGURE 126

[illegible]

127/330

FIGURE 127

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCK
CSFNQKPRAPGDEEAQVENLITANATEPQKQRTEVQPSGGSWNLRLLEPLDANVDA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 128

AAACTTGACGCC**ATG**AAGATCCCGGTCCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT
CCTCTGCCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG
CGTCACGACCCGAGGCCTTTAACACCCCGTTCCTGAACATCGACAAATTGCGATCTGCGTTT
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAAGGAACTTCCTTT
CCTCAACTGGGATGCCTTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAG**T**
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT
CTTTCCTGCCTCAGGAACTCCAATAAAACATTTTCCATCCAAA

129/330

FIGURE 129

MKIPVLPVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSFAKDE
FLNWHALFESIKRKLPFLNWDAFPKLKGLRSATPDAQ

0996041-14036560

FIGURE 130

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCC**ATG**GGGGTGGAGATTGC
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTT
TTCTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC
TTTGCCGGCCACTCA**TGAG**AGTGTTTTTGTGTAAAGTATTTTTTAGAATACTGTTGACTTCT
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCA
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT
AGGTTCAAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTTTATC
TTCACCTATTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTT
TATCCCTGGGTGCCCCTGACACATTTATGTAGTGATCCCACAAATGTGATTGTTAATTTAAA
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT
AAAAGTTTTGAGTATATTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

131/330

FIGURE 131

MGVEIAFASVILTCLSLLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

090904 1430
"T" 405050

FIGURE 132

GGGGAATCTG CAGTAGGTCTGCCGGCGAATGAGTGGTGGGCTAGCTCGCCGCTTCCGGCTCTG
GCTGCTGTTGTTCTCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGA
AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC
TGCAGCTGCTACCATGGTGT CATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG
GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA
GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTGAGCACTTTATT
TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA
GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC
ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT
CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG
GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG
ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACCTTGTTGATGCAGAATACACCAAAAAC
CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT
GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA
AACACCTCTTCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC
TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCA
AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA
GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG
AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT
TCCCAAAATGTTGAAAACCTGAACTATAGTAGTCATCATAGGACCATAGTCCTCTTTGTGGCA
ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA
TCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAGAGCAACTCTTGAGA
AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA
CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC
ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA
TGTGATGATGCCCTTTGTCCCATTTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA
CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG
AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG
TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTA AAACCATAAACTCTGTTACTCAG
GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT
TCAGGTTCCCTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

MEWWASSPLRLWLLLFLLP
SAQGRQKESGSKWKVFID
QINRSLENYEP
CSSQNCSCYHGVIE
EDLTPFRGGISRKMAE
VVRRKLGTHYQITKNRL
YREND
CMFPSRCSGVEHFILE
VIGRLPD
MEMVINVRDYPQVPK
WMEPAIPVFSFSKTSE
YHDIMYPAWTFWEGGP
AVWPPIYPTGLGRWDL
FREDL
VRSAAQWPWKKNSTAY
FRGSRTSPERDPLILL
SRKNPKLVDAEYTKN
QAWKSMKDT
LGKPA
AKDVHLVDHCKYKYL
FNFRGVAASFRFKHL
FLCGSLVFHVGDEWLE
FFYPQLKPWVH
YIPVK
TDLSNVQELLQFVKAN
DDVAQEIAERGSQFIR
NHLQMD
DITCYWENLLSEYSK
FLSY
NVTRR
KGYDQIIPKMLKTEL

FIGURE 134

CACCCCTCCATTTTCTCGCCATGGCCCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTT
TCATCCTGGCCTTTGGCACCGGAGTGGAGTTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT
GGAGGGATCCCGGAGTCTGGTGGTCCGGATGCCCGCCAGGGATGGCTGGCTGCCCTGCAGGA
CCGCAGCATCCTTGCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC
ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCTTTCAG
AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCACTGGTGATGCGGTACTGGGAGCC
CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC
TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT
GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC
TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCCACCTGCGCCACCCAGTGTGTG
TGGAGCTGCTGACAGTGCTGTGGGTGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT
TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT
CCGGGCCCAGCTACAAAGAAAACCTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAGT
GAGGAGCTCACTCTGGTTACAAGCCCTGTTCTTCCTCTCCCACTGAATTCTAAATCCTTAAC
ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCTT
CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA
CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC
TTCCTGTTTGTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC
CTGACCACTCCCCTGGCACTGTTACTTGCCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT
GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA
GGCCCCAACCTTGCCTCACCCTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT
GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC
CCTGCTCTCCCGAGGAAGATGCTCTGCAGGAAAATAAAAGTCAGCCTTTTTTCTAAAAAAA

FIGURE 135

MAPALLLIPAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAALQDRSILAP
LAWDLGLLLLFVQGHSMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV
LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP
RALRLFSLRHPVCVELLTVLWVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR
KLHLLSRPQDGEAE

Signal sequence:

amino acids 1-13

Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

FIGURE 136

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA
 AGAAATTGCCAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGT
 TTTAATTAAGCATGGAATACAGAAAACAACAAAAAACTTAAGCTTTAATTTTCATCTGGAATT
 CCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTA
 TCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTAAGAACCTTCGGCTCGCGTGCTTCTG
 AGCTGCTGTGGATGGCCTCGGCTCTCTGGAAGTGTCTTCCGAGTAGGATGTCACCTGAGATCC
 CTCAAATGGAGCCTCCTGCTGCTGTCACCTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT
 TCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTT
 ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTT
 CTGGTCATTCTGGTGACCTCCCACCCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTAC
 TTGGGGTGAAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTTCTTATTAGGCCAAG
 AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT
 GACATAATCCGACAAGATTTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTTATGGC
 ATTCAGGTGGGTAACTGAGTTTTGCCCCAATGCCAAGTACGTAATGAAGACAGACACTGATG
 TTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT
 TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT
 TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATATAA
 TGTCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAAACCCATCAAGTTT
 GAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAAAAGTGAACATTCATATTCCAGAAGA
 CACAAATCTTTTCTTTCTATATAGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTG
 CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTTGGCAGGTCATGCTAAGGAACACC
 ACATGCCATTATTAACCTTCACATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA
 GTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG
 AACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC
 AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA
 GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG
 TGTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTATTG
 AACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA
 TGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT
 TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATTATTATTTAAATTA
 CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT
 CATCTTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC
 TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAAT
 ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAA

FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQD
FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFELLGQEA EK
EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFIN
TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD
LVPRIYEMMGHVKPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG
FSSKEIITFWQVMLRNTTCHY

FIGURE 138

CCTCTGTCCACTGCTTTTCGTGAAGACAAGATGAAGTTTCACAATTGTCTTTGCTGGACTTCTT
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAAATGGCTTTGCTGCAACC
AGACTCTTTCAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT
TCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC
CTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA
AAAAACATTGCAAACATGTGTCTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC
AAGCCTGTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT
CCTTCTGTGGAGACACGGTGGAGAACTTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT
TTTCTCTACTAGTTATGTTTTGATTTCTTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAAA

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNGWDSWNS
IWDYGNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN
PNKVDDLSKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

FIGURE 140

CATTCTGAACTAATCGTGTGAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATA
 TTAACTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG
 CTCATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCCTATTGCTTACTGATTTTTT
 TGAGTTAAGAGTTGTTATATGCTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAA
 GAATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGCTGGTTTGTGTACATG
 CAAGCTTATAGTTGAAATATTTTTTCAGGAATTAC**ATGA**ATGACAGTCTTCGAACCAATGTGT
 TTGTTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT
 CAGATTCCGTTGCCAACTCGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAAT
 CCAGGAAATCTGCATAGAAACACTTAGGCTTTATACCAGAAAAAGCCAACTATGAATTAC
 TGGAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCAAATTTAAAGCAAAGGGA
 TTGAATCCGGATGGAAGCTCCAGCCCTTCAACCCCTGGGTGGATTTTCTCCAGCCTCCAAGCC
 ATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG
 TCAAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAGAAAA
 GACAGCAAGAGAAGTAGAAATAGCAGAAGTGCAAGTCGATCGAGGTCAAGAACACGATCACG
 TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA
 GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT
 GGTCTCCTCACCTTAAGGCCAAGCATAACCAGAGATGATTTAAAAAGTTCAAACAGACATGG
 TCATAAAAGGAAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG
 CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTT
 GAGAGGTCCCATAAAAGCAAGCACCATGGTGGCAGTCGCTCAGGACATGGCAGGCACAGGCG
CTGACTTTCTCTTCCTTTGAGCCTGCATCAGTTCTTGGTTTTGCCTATCTACAGTGTGATGT
 ATGGACTCAATCAAAAACATTAAACGCAAAGTATTAGGATTTGATTTCTTGAAACCCTCTA
 GGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTGCACATT
 AAAATGCCCTAGCAGTATCTAATTTAAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT
 TGTGTATTGTTTATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTT
 ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAAT
 ACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTTACAAGGAAATAAAATACAAAT
 CTTGTTTTTTCTAAAAAAAAAAAAAAAAAAGT

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEK RKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAE EK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHK RKKSRSRSQ
SKSRDHSDAAKKHRHERGHHDRDRERSRSFERSHKSKHHGGSRSGHGRHRR

FIGURE 142

TGGGGATAAAGGAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA
TTTTTTGAANNTATTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT
CATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTNNTTATTGCTTACTGATTTTTTTG
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAGA
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGTGGTTTGTTTACATGCA
AGCTTATAGTTGAAATATTTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT
GTTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCATCTACCTTGCAGNTAGAGCACTTCA
GATTCCGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCC
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAGCCAACTATGAATTACTG
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTAAAAGCAAAGGGATT
GAATCCGGATGGAAGTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCC

FIGURE 143

GGCACGAGGCCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTC**ATGGC**
GGTCCTCGGAGTACAGCTGGTGGTGAACCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC
CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCGATAACAAGCACCCG
TCTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCCAGAGGCAGGAAAGAGCGGTG
GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG
AGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG
TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACTACAT
GCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGTGCCTGCTCACGGTGACCTTCT
CCATCAAGATGTTCCCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC
TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGCCATGCTGGTGCAAGTGGTGCG
GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC
CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG
GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC
CCAGACCCACCGGGACGCACTGACCATGTGCGAGGACAGACCCATGCTGCAGTTCCTCCTGC
ACACCAGCTTCCTGTCTCCCCTGTTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC
TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA
CTCTGGGCGCCTCTGGTTGCTGGTGGTGCTGTGCCTGCTGCGGCTGGCGGTGACCCGGCCCC
ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC
CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT
GAGCTTGACGTACCTGACGCCGCTCATCCTCACCTCAACTGCACACTTCTGCTCAAGACGC
TGGGAGGCTATTCTGGGGCCTGGGCCCAGCTCCTCTACTATCCCCGACCCATCCTCAGCC
AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG
GGCCCTGGGTGGCCTGCTTACTCCCCTCTTCCTCCGTGGCGTCCTGGCCTACCTCATCTGGT
GGACGGCTGCCTGCCAGCTGCTCGCCAGCCTTTTTCGGCCTCTACTTCCACCAGCACTTGGCA
GGCTCCT**TAG**CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCTGGGGCAGCGGGACA
CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCAGCTGCAAGGTGGGGCCGGACTCCCC
GGCGTTCCTTCACCACAGTGCCCTGACCCGCGGCCCCCCTTGGACGCCGAGTTTCTGCCTCA
GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG
TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC
GATTTTTTAA

FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEELRALAGKPRPRGRKE
RWANGLSEEEKPLSVPRDAPFQLETCPLETTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY
YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV
VREETLELGLEPGLASMTQNLPLELLKKQGWDWALPVAKLAIRVGLAVVGSVLGAFITFPGLR
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT
VVSLQYLTPILITLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAAARI
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLAAS

FIGURE 145

CGTTNGCACGCGTCAATGGCGGTCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTT
TGTTCCGATAACAAGCACCCGTNTTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCC
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTCCCTGACAGTGACACGGCTGTACTTCAGC
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC
CATGCTGGTGCAAGCG

FIGURE 146

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTT
AATCTGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGC
TTGGCTGGTTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGA**ATGA**AGG
CGCTTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTG
TATTCAGAACTCTGTAAAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTG
TCCAGACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAA
TGACAGACGAGCCTGGCCTAGACAACCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCCA
GTGGACTCTGGCCGGAGCAACCGAAGTGGGACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA
AAAAATAAATCGAGCTTTGAGTGTTCTTCGAAGGACAAAGACGGGAGTGCAGTTGCCAACCATGCCGACCAGG
GCAGGGAAAATTCTGAAAACACCACTGCCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAA
ATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC
CCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGGTGATCGCCAGAGACGGCCGGCTACTGCCAG
GAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCCTGCGG
CAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCGCGCAGCAGGAACAATGGACAGGCCCC
GGATGCCTACAGACCCCGAGATGACAGCTTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAA
TAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGA
CATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCCAGAAAG
TGCGGCTCATCTGATTGAGGCCAGTGAAAGACGTGTTACCTCGTCTGTCCCGCCAGGTTCCGCAGCGGAGCC
CTGACATCTTTGAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACT
CCCCAGCCCCTCCATCCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCT
CGGCATGACCGTTCGAGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTGAGCCCG
GAGGAGTCATAAGCAGAGATGGAAGAATAAAAAACAGGTGACATTTTGTGTAATGTGGATGGGGTCGAACTGACA
GAGGTGAGCCGGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGT
CAAAGAGTATGAGCCCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCA
GTGACTGGTCCCCATCCTGGGTGATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTA
CGAAGAAACACAGCTGGAAGTCTGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCTTT
TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTTCTTG
CTGTCAATGGTAGAAGTACATCAGGAATGATACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGA
ATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTTTTAT**TAGA**ATCAATGATGGGTGAGAGGAAAACAGAAAAA
TCACAAATAGGCTAAGAAGTTGAAACACTATATTTATCTTGTGAGTTTTTATATTTAAAGAAAGAATACATTGT
AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAATTA
AACTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTACAACATTGTTTATATTTTTTCTATTCAAT
AAAAAGCCCTAAAAACAATAAATGATTGATTTGTATACCCCACTGAATTCAGCTGATTTAAATTTAAAAATTT
GGTATATGCTGAAGTCTGCCAAGGTACATTATGGCCATTTTTAATTTACAGCTAAAATATTTTTTAAAAATGCA
TTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTTCAGAAGTTAAA

147/330

FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT
APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSEGRSNRTRARPFERSTIRSR
FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDP
SELSIRLVGGSETPLVHIIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLL
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVSRQVRQRS
PDI FQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE
WDLPIYVISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV
KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV
GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRC GDILLAVNGRSTSGMIHACLARLLKELKGRI
TLTIVSWPGTFL

009904.1.40660

FIGURE 148

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCT
GCACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAA**ATG**GTGCTGACCATCT
TTGGGATACAATCTCATGGATACGAGGTTTTTAACATCATCAGCCCAAGCAACAATGGTGGC
AATG TTCAGGAGACAGTGACAATTGATAATGAAAAAATACCGCCATCGTTAACATCCATGC
AGGATCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGC
TCTCCCGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAAT
CTCCAATGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTG
GGTCAAGTACAACCCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTAC
CCATTGAGAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACAT
AATGTCGGTGCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTG
TGCAGACATTCATGTT**TAG**GATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCC
TTGGTTTACACTCAAAGTCAAATTAAATTCTTTCCCAATGCCCCAACTAATTTTGAGATTC
AGTCAGAAAATATAAATGCTGTATTTATA

149/330

FIGURE 149

MKILVAFLVVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNQLQWYIYEKQALDNMFSNKYTWVKYNPLE
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

149/330

FIGURE 150

GGCACGAGCCAGGAACTAGGAGGTTCTCACTGCCCAGCAGAGGCCCTACACCCACCGAGGC
ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG
CCAGTGTCTGAGCTGCGGGCCAACCTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGA
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA
GCCCCATCCTTGCCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG
GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAGA**ATGAACC
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTTCGTATTTGGA
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVPKGRWVLITCCAPQPPPPITYSL
CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK
PVSELNANFTLQDRGAGPRVEMICQASSGSPITNSLIGKDGQVHLQQRPCHRQPANFSFLP
SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLES PILALPLYRSTRRLSEEEFG
GFRIGNGEVRGRKAAAM

Signal Peptide:

amino acids 1-18

N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

FIGURE 152

GGTCCTTAATGGCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG
CTGTCCGGCTGGTCCC GGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT
CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA
CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA
CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT
TACAGAGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCCTGC
AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT
TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC
TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCTTCCATT
ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC
CTGGAGCCAAGTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACCTCAGGGCCAC
AGCCACCACCCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA
TCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCTGATACCAAAGGCTCCTGTGAGCACG
GTCTTGATCAAACCTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT
GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCCTT
TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC
TGATGGAATTCCTGCACTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTC
TCTTTTTGTTTGGAAAATCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATT
GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG
AATTTTTTAAATTATTTAATAAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT
TGTTCTGTACTGATATTTAATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 153

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFL
HYDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

FIGURE 154

GGGAAAGCCATTTGAAAACCCATCTATACAACTATATATTTTCATTTCTGCTGCTAGCTG
CCTTGGGCCTCACAATTTTCATTCTGTTTTCTGACTTTCAAGTTATATACCGTGGA**ATG**GAG
TTGATCCCAACCATAACATCGTGGAGGGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTG
GTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAATCATGAACTCTGGCTGTTGATCAAAA
GAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAGAAGCTAGCAGAAGAC
TCAACCTGGCCTCCCATAAACAGGACAGATTATTCAGGTGATGGCAAAAATGGATTCTACAT
CAACGGAGGCTATGAAAGCCATGAACAGATTCCAAAAGAAAACCTCAAATTGGGAGGCCAAC
CCACAGAACAGCATTTCTGGGCCAGGCTG**TAA**TCAGAATTGTCGTCGTACATGCTCAACAGC
ATTGCTTTTTTCCCCAAAATTAACACATTGTGGAGAAGTGATGATACTCTCCCCTTACCTTT
CCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATTAAACCTTGCAGCAAGGGACC
TTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATTCCTGTAT
CATCCTTTTCAATAAACTGTATTCATTTTGAAAAAAAAAAAAAAAAAAAAA

155/330

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHELWLLIKREFGFYSKSQYRTWQKKLA
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGQPTEQHFWARL

099304.11.01
"TTP" 408660

GTTCTCCTTTCCGAGGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACCC**ATGAAG**
 CTCTTGTGGCAGGTAACCTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTTCGTCTA
 CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCTCAGCCGGGGCCCC
 AGAACTGCCCCCTCCGTTTGCTCGTGCAGTAACCAGTTCAGCAAGGTGGTGTGCACGCGCCGG
 GGCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA
 CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTGTCAGT
 TGGGCAGGAACTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAAC
 ACCCTGGAGCTGTTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC
 CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCCATCGAAAGCATCCCCTCTTACGCCTTCA
 ACCGGGTGCCCTCCCTCATGCGCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT
 GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA
 AGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACT
 TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCTATG
 AACTCACAGGTGAGCCTGATTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAAC
 CAACTTGGCCCAATAACCTCTCTTCTTTGCCCCATGACCTCTTTACCCCGCTGAGGTACC
 TGGTGGAGTTGCATCTACACCACAACCCTTGGAACCTGTGATTGTGACATTCTGTGGCTAGCC
 TGGTGGCTTCGAGAGTATATAACCCACCAATTCCACCTGCTGTGGCCGCTGTCTATGCTCCCAT
 GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT
 GCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTTCGGATGGCAGAACTTAAGTGT
 CGGACTCCCCCTATGTCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC
 CTCCCGCCACCCAAGGATCTCTGTCTCAACGACGGCACCTTGAACTTTTCCACGTGCTGC
 TTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG
 GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACCTACAGCTTCTTACCACAGT
 AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACCGGAAAGTACAAGCCTGTTCTTA
 CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACC
 CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT
 GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG
 CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC
 ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCAGCAGCAACATCCGC
 AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCACAATTCA
 ATGACCATATTAACCTACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAAACAGC
 CTGGGGAACTCTCTGCACCCACAGTCACCCTATCTCTGAACCTTATATAATTTCAGACCCA
 TACCAAGGACAAGGTACAGGAACTCAAATA**TGACT**CCCCCTCCCCCAAAAACTTATAAAAT
 GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA
 TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAACAGATTATATTAAAATTTAAGA
 CAAAAAGTCAAAACA

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGI PSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLEELEM SGNH FPEIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHPWNCDCDILW
LAWWLREYIPTNSTCCGRCHAPMHRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGTLNF SHVLLSDTGVYTCMVTNVAGNSN
ASAYLNVSTAELNTSNYSFFT T VTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTTDKMQTSLDEVMTTKIIIGCFVAVTLLAAAMLIVFYKLKRHQQRS
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNSLHPTVTTISEPYIIQTHTKDKVQETQI

FIGURE 158

CGCTCGGGCACCAGCCGCGGCAAGGATGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCC
 TTCTCATCTCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATG
 TGTGCGGGAGTGCTGTGAATATGATCAGATTGAGTGCGTCTGCCCCGAAAGAGGGAAGTCGTGGGTTATACCAT
 CCCTTGCTGCAGGAATGAGGAGAATGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAAC
 GCAAGAGCTGCCGAAATGGCTCATGGGGGGGTACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAG
 TGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTT
 GTTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGCTAAACCTGGGTTTGTATCCAACATAA
 GATTTGTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTTCGTGATGGAGACAAC
 CGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAGGATCCTCACT
 CCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGCAT
 GCTCCTCATCCCTTGTTTTCCATGACGGCACGTGCGTCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTG
 GCAGGCTATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAAACTGCTCAGACCCTGGGGGGCCAGTCAA
 TGGGTACCAGAAAATAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATTGGCACCCTGGTGTCTT
 TCTTTTGTAAACACTCCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGG
 AAACAGCCCCTGTCATAAAAGCCTGCCGAGAACCAAGATTTTCCAGACCTGGTGAGAAGGAGAGTTCTTCCGAT
 GCAGGTTCAAGTCAAGGGAGACACCATTACACCAGCTATATCTCAGCGGCCCTTCCAGCAAGCAGAAAACGAGAGTG
 CCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGGATACCAACATCTGCATACCCAGCTCCAG
 TATGAGTGCATCTCACCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGTCTGAGGACTGGGAAGTGGAG
 TGGGCGGGCACCATCCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGACCAAGGGTTGC
 GCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGG
 TTCCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTTACTGACCTGGG
 GAAGGTCAACATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTCTACCGGGATGATGACCGGGATG
 AGAAGACCATCCAGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCTGCTTGATGCT
 GACATCGCCATCCTGAAGCTCCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAG
 TCGGGATCTCAGCACTTCCTTCCAGGAGTCCACATCACTGTGGCTGGCTGGAATGTCCTGGCAGACGTGAGGA
 GCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGGTGGTCACTGTGGTGGACTCGCTGCTGTGTGAGGAGCAG
 CATGAGGACCATGGCATCCAGTGAGTGCTACTGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCTTC
 TGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCTTCCCGGGACGAGCATCTCCTGAGCCACGCT
 GGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCACTGCCTTACCAAG
 GTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAATGGAACCATGCTCATGCACTCCTTGAGAAGTGTTTC
 TGTATATCCGTCTGTACGTGTGTCTTGCCTGAAGCAGTGTGGGCTGAAAGTGTGATTGGCCTGTGAACTTGG
 CTGTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGAT
 GCCGCTCCACTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGAAGAAGTAAGTTTCTTCAAAGAAGACC
 ATATACAAAACCTCTCCACTCCACTGACCTGGTGGTCTTCCCCAACTTTTCAAGTTATACGAATGCCATCAGCTTG
 ACCAGGGAAGATCTGGGCTTCATGAGGCCCTTTTGGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCC
 CAGGGCAGCAGAGCTGGGATGTGGTGCATGCCTTTGTGTACATGGCCACAGTACAGTCTGGTCTTTTCTCTCC
 CCATCTCTTGTACACATTTTAAATAAAATAAGGGTTGGCTTCTGAACACAAAAAAAAAAAAAAAAAAAAAAAAA
 AA

FIGURE 159

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDEFYVKGFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLFSHSDGSKNFDGFAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPLH
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVAAHCVTDLGKVTMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTS FQESHITVAGWNVLADVRS PGFKNDTLRSG
VVSVDSSLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA
 AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAAC**ATGGGC**
 TTCAACCTGACTTTCCACCTTTCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT
 GACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCTAAAG
 CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAAACTCTGACT
 AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACCTGTCCTTCTGTGTCTCCTTACCTCAG
 AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC
 CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC
 CTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT
 CCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT
 TTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC
 TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA
 GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG
 GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCT
 AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG
 AATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG
 ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACGAGTCTGG
 AGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA
 TATCAACATCACAGTGGATTTCTGGTTTGGTGCA**TGAC**CCCTGGATCTTTTGGTGATGTTTGG
 AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA
 AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTTGTATTTTCTTAGCAGAGCT
 CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGATCATG
 AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT
 AAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATAT
 TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT
 CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTCAATTTATCCTGTACAATCATCT
 GTGAAGTGGTGGTGTGAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA
 GGACACAGTGAACCTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAG
 CAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT
 TCCAGTGATGCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAAATGA
 TTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAATAATA
 TGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

161/330

FIGURE 161

MGFNLT FHLSYKFRL LLLLLTLCLTVVGWATS NYFVGAIQE I PKAKEFMAN FHKTLILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAE GKKFNRAKLLNVGYLEALKEEN
WDCFI FHDVDLVPENDFNLYKÇEEHPKHLVVGRNSTGYRLRYSGYFGGV TALSREQFFKVNG
FSNNYWG WGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWR TDGLSSCSYKLVSV EHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

T05T04T03T02T01

FIGURE 162

CGTGGGCCGGGGTTCGCGCAGCGGGCTGTGGGCGCGCCCGGAGGAGCGACCGCCGCAGTTCTC
GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCA**ATG**
GCCCAGGCAGTGTGGTCGCGCCTCGGGCCGCATCCTCTGGCTTGCCTGCCTCCTGCCCTGGGC
CCCGGCAGGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA
CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG
CCCGCTGACGCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGCTTACTGGCAA
GATGGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTTCGGCCACGTGCCCCGGGAATTCCCGG
TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG
GTCCTCCCCATCACAGAGTTCCTCGTGGGGGACCTTGTTGTCACCAGAACAACCTTCCCTACC
CTGGCCAGCTCCTATCTCACTAAGACCGTCTGAAAGTCTCCTTCCTCCTCCACGACCCGA
GCAACTTCCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACGGGACCCAGATG
GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT
CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGA
CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCCTTCGAGGCATCCAAGTGTGGGG
CCCACCCTAATTCAGACCTTCCAAAAGATGACCGTGACCTTGAACCTCCTGGGGAGCCCTCC
TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC
CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCCTGGGGACTAC
TGCTTCAGCATCCGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT
GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA
TGTTGGCCTTCATCATGTACATGACCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG
AACCCGGAGCCACCCTCTGGGGTCAGGTGCTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT
GGAGACTCCATCTGAGTACCTGGAAATTGTTCTGTGAGAACCACGGGCTGCTCCCGCCCCTCT
ATAAGTCTGTCAAAACTTACACCGTGT**AG**CACTCCCCCTCCCCACCCCATCTCAGTGTTAA
CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCAATT
TGCGTGGGGCTGTTGGCCTGGATCATCCATCCATCTGTACAGTTCAGCCACTGCCACAAGCC
CCTCCCTCTCTGTACCCCTGACCCAGCCATTACCCCATCTGTACAGTCCAGCCACTGACA
TAAGCCCCACTCGGTTACCACCCCTTGACCCCTACCTTTGAAGAGGCTTCGTGCAGGACT
TTGATGCTTGGGGTGTTCGCTGTTGACTCCTAGGTGGGCCTGGCTGCCCCTGCCCATTCCT
CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCTCCCCAGACAGCCCTAC
CTGTGCCAGAGAGCTAGAAAGAAGGTCATAAAGGGTTAAAAATCCATAACTAAAGGTTGTAC
ACATAGATGGGCACACTCACAGAGAGAAGTGTGCATGTACACACACCACACACACACACA
CACACACACACAGAAATATAAACACATGCGTCACATGGGCATTTTCAAGATGATCAGCTCTGTA
TCTGGTTAAGTCGTTGCTGGGATGCACCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA
AGCAGCCCTGACAGGTTCTGGGCCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC
GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCCCTGGATGGGGGGCAGGACT
AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG
AACTTTCACTGAGGAAAAGGCCTTGCAGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGG
TGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA
GATCGAGACCACCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAAATACAAAAAGTT
AGCCGGGCGTGGTGGTGGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG
GTGCGAACCCGGGAGGCGGAGCTTGCACTGAGCCAGATGGCGCCACTGCACTCCAGCCTGA
GTGACAGAGCGAGACTCTGTCTCCA

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWSSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTNLFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMC CGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

FIGURE 164

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATG**GCACTGAGCTCCCAGATCT
GGGCCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCA
CAACAGACGGGACAACCTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTG
GATGCCCATGTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCT
GCGGCTGCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAGA**ACCTACCTGCCCTG
CCCCGTCCCCTCCCTTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAA
TGGCTGGTTCTTTGTTTTCCAAA
AAA

165/330

FIGURE 165

MALSSQIWAACILLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH
FPICIFCCGCCHRSKCGMCCKT

165/330

FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC
 CTGGATCTTCCACCATGTTTCCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC
 ATCTCCCTGACTGTCTCTTCACCCTCCTTCTCGTTTTTCATCATAGTGCCAGCCATTTTTTGG
 AGTCTCCTTTTGGTATCCGCAAACCTCTACATGAAAAGTCTGTTAAAAATCTTTGCGTGGGCTA
 CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC
 GGAATCATTGCAAAGGATCCCACCTTCACTAGAAAGAAGAGATCAAAGAGATTTCGTCTGAAGTGG
 TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCC
 GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG
 GAGTCCTGGAACCTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC
 GGTCTGTGGGGGTAGGAGTGCTGATTCGGTACTGCTTTCTGCTGCCGCTCAGGATAGCAC
 TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACCTGTGGTGGGATACTTGCCAAAT
 GGGAGGTTTAAAGGAATTCATGAGTAAACATGTTCACTTAATGTGTTACCGGATCTGCGTGCG
 AGCGCTGACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT
 GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC
 ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTGAGAGAGCCATGGTGAAGGCCTG
 CCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA
 CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCCAGAAGGAACCTGCATC
 AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTGAAATTGGAGCCACAGTTTACCC
 TGTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA
 TGGTGACGTACCTGCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG
 CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC
 CATTGCCAGGCAGGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCCTGAAGAGGGAGAAGG
 TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC
 AAGGACAGGAGCCGCTCCTGAGCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAA
 CGGGCTCAGAGCTGGAGTTGCCGCCGCCGCCCCCACTGCTGTGTCTTTCCAGACTCCAGGG
 CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT
 GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA
 CGAGATGCCTTGTTTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACCTCCCA
 CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT
 GCGGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGCGGCCACCCG
 CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGC
 CTTGGAGCTCTGCAGACATGATAGGAAGGAAACTGTCATCTGCAGGGGCTTTCAGCAAATG
 AAGGGTTAGATTTTTTATGCTGCTGCTGATGGGGTTACTAAAGGGAGGGGAAGAGGCCAGGTG
 GGCCGCTGACTGGGCCATGGGGAGAACGTGTGTTTCGTACTCCAGGCTAACCCTGAACTCCCC
 ATGTGATGCGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGG
 GAATGGTGGTGATTCCCTACCTCACAGGGCTGTTGTGGGGATTAAAGTGCTGCGGGTGAGTGA
 AGGACACATCACGTTCAAGTACAGGCCCAAAAACGGGGCACGGCAGGCCTGAG
 CTCAGAGCTGCTGCACTGGGCTTTGGATTTGTTCTTGTGAGTAAATAAACTGGCTGGTGAA
 TGA

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIVPAIFGVSFGRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRSKSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSAAEELESWNLLSRTNYNFQYISLRLTVLWGLGVLIRYCFLLPLRIALAF
TGISLLVVGTTVVGYLPNGRFKEFMSKHVHLMCYRICVRALTAIITYHDRENRPNGGICVAN
H TSPIDVILASDGYIAMVGQVHGGMLMGVIQRAMVKACPHVWFERSEVKDRHLVAKRLTEHV
QDKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMV
TYLLRMMTSWAIVCSVWYLPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVK
DTFKEEQKLYSKMIVGNHKDRSRS

GCCCCTCGAAACACGAGACTCCAGCACCTCTGGTCCCCTCACCCGGACCCCTGGCCCTCA
 CGTCTCCTCCAGGGATGGCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC
 ACCTGGCAGGCCCAGGCTGTTCCCACCATCCTGCCCTGGGCCTGGCTCCAGACACCTTTGA
 CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG
 AAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCAGGAGACCTGGGAGGAC
 AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT
 CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG
 GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG
 GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCG
 AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT
 TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT
 GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCTG
 GAAGACTCTGCTCTTGGCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCCTGAAAGTCCA
 ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG
 CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG
 ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTATGTGATGGGGACTTCCT
 GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTGCAATGTGGAGACA
 TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

169/330

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEKAAPLLKEEMAHH
ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSREL
YMRHFPPKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS
LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

169/330

FIGURE 170

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAAT**ATG**GCTGGTTCCCCAACATGCCTCA
CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG
GTCGGTTCGGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC
TATTGTCTGGACCTTCAACACAACCCCTCTTGTACCATACAGCCAGAAGGGGGCACTATCA
TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG
CTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT
CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG
TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG
GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC
CCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT
GCGTTGCCAGGAACCCTGTGAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT
GAAGGTGCTGCTGATGACCCAGATTCTCCTCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCCT
CCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAG
AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGGAAACTCCTAACATATGCCCCCAT
TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA
TCCAGCAAATACGGTTTACTCCACTGTGGAATACCGAAAAAGATGGAAAATCCCCACTCAC
TGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATC**TAG**ACAGCAGTG
CACTCCCCTAAGTCTCTGCTCA

171/330

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSPKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEEYIEEKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

090304.1.1.1.1

FIGURE 172

CTGGTTCCTCAACATGCCTCACCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC
TCTGGACCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC
AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA
GATGGAGGCTACTCCCTGAAGCTCAGCAAAGTGAAGAAGAATGACTCAGGGATCTACTATGT
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCTCCATGGTCCTCCT
GTGTCTCCTGTTGGTGCCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAA
ACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA
TAGAACCAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG
AATGTTATCTAGACAGCAGTGCCTCCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAAA**ATG**AAGATGCTGCTGCT
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGAGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCCTCTTCCTGTTCAATAAATGATTACCCTTGCACTTAA

174/330

FIGURE 174

MKMLLLLCLGLTLVCVHAEESSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ
IHVLENSLVCLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI
NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

Downloaded from www.oxfordjournals.org/

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAA
TGGATTTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTG
TCAGCTTAGTTGAGGAAGACCAATTTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTTC
CCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAG
AAAAAGAGCGTGCTGCAACAACAGAACTGGAATGTTTCTTTTCATCATTTTTTCAGTGTGATCA
CAGTCATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCCTCTC
ATGTGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGA
CATTCATCCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTG
GTTTCAATAAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCAC
TTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTTAGGTCTATTGCT
TGTTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGCTGTC
TGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTGTAGTTTAATGGGAATAAAATGTAAGTA
TCAGTAGTTTGAAAAAAAAAAAA

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPI SCFEWWFPGI IGAGLMA
IPATMSTLTARKRACCNNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLM C NSPNSNA
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDT MASGWRASSFHFDSEENKHRL
IHFSVFLGLLL VGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGAGGCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTTGAAAAAGTCCTGGTGGAAATAGTGAA
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC
CTGATCTTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

178/330

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV
KHCTDQISFKKRLSLKKSWWK

178/330

FIGURE 179

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG
GAGCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTC
GAACTGTGAC**ATG**GAGAGAGTGAACCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCAC
TCATCACTCCAGGCTCTGCCACTACTTGCT**TGAG**CACAGGACTGGCCTCCAGGGATGGCCTGA
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC
TCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAAACTCG
CCCCACCACCCCCTCA

180/330

FIGURE 180

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCK
YKSSQKQHSPVPEKAIPLITPGSATTC

099904.4.44501
T0577.408550

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCC
 TGGCTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC
 CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACTGCCGCCGGCTCC
 AGTGTTTCCACAGCCCCCAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCCT
 ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTGCGCCACCTATTCACAGGGCTTTACGGT
 ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA
 CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGG
 CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC
 GCCCGCCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAACA
 TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG
 CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG
 TCAGGAGAGGCCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA
 GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGC
 TTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG
 CACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG
 ATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT
 ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC
 CTGGGTCCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC
 AAGAGCTTCTGAAGGACCGGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC
 TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG
 ATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC
 TCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGAC
 CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCTTTCTC
 CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGG
 CGTTGATGCTGCTGCACTTCGGGTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAA
 TTGATCATGCGCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA
GTGACTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAA

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGEKILLSGGDKWSRHRRLTPAFHFENILKSYITIFNKSANIMLDKWQHMLASEGS
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLDFIDVLLLSKDEDG
KALSDEDI RAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

FIGURE 183

CAACAGAAGCCAAGAAGGAAGCCGTCTATCTTGTGGCGATCATGTATAAGCTGGCCTCCTGC
TGT TTGCTTTTCACAGGATTCTTAAATCCTCTCTTATCTCTTCCTCTCCTTGACTCCAGGGA
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAACTCCGGAGGAGCTAGAAA
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG
AAAGCAGACTCAAGTACCAACATTTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATACA
AGAAACGTGAGACTCCTGATTGCTTCTGGAAATACTGTGTCTGAAGTGAAATAAGCATCTGT
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAC
AGTGTGGAGAAAACTAGGCAAACCTACACCCTGTTCAATTGTTACCTGGAAAATAAATCCTCT
ATGTTTTGCACAAAAAAAAAAAAAAAAA

FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDSREISFQLSAPHEDARLTPEELERASLLQILPEMLGA
ERGDILRKADSSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETDCFWKYCV

FIGURE 185

GAACATTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT
GGGGTTGCTGGTTTAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA
CCACCTCCGCCAGGAACTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCCATGCCCTCCCC
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCA
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT
CAGGGGTTCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAAGTTTCTTCAGGACATCCTCTGG
GAAGAGGCCAAAGAGGCCCCAGCCGACAAGTGATCGCCACAAAGCCTTACTCACCTCTCTCT
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA
CAAGCTCAGGAGGCGAATAAATGTTCAAACCTGTA

FIGURE 186

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG
GQAEGAEDLEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC
GTGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAATATGAACACGTGGCTGCTGT
TCCTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG
CTCCTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT
GAGTATGTCCCCACCCTAAGCCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC
TTACACCTCTACTTGAGTATGTCCCTAACCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT
GTCCCCCGTGTGCGCATGTGTTTCAGGGTCAGCCTCTCCAGAAGTGAGATCATGGACAAAAA
GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCAGCAAGAAGCTGAAC
TCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCTTGAAGTAACAAGTTTAAAATGTTCAGA
GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT
CCAAAAACACAAGTAGAAATTCTAACAATGAAATATATTACAGGCAGGTCACCCACTAACCA
AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC
AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT
AACAAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTTGGTCATGGTGGGCCAGCTGCAAAGCG
TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTCTTCTGTGGAC
ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA
ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTGAGGCTCTGGGCTCACCTCCATCTCAGA
GCATCCCCTGCCTGCAGTTGTGGCAAGAACGCCAGCTCAGAATGAACACACCCACCAAGA
GCCTCCTTGTTTCATAACCACAGGTTACCCTACAAACCACTGTCCCACACAACCCTGGGGAT
GTTTTTAAACACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA
ATTTTTTTTTTAATGAAAGTGCAATGAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

188/330

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLDFLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW
VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ
AQQEAELTPRPAGVVPGA

Downloaded from www.jstor.org

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCCTTCCAGACAAGCTGCAAGACGCTGACCATGGCCAAAG
ATGGAGCTCTCGAAGGCCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG
TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG
GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACAACCTGGGAGACTGGGGATGA
CCGGTTCCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAG
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA
GCGGTTGATGGAGAAGGCTTCCCTCCCCCTCCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA
TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGTTCCCTCCTGCT
ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACCTGAGCGCCTTTGCTG
CTGTTTCCTCTGTCTCTGTGAGTCTCCTGGGGATGGTGGCCACATGATGTATTCACAAGTC
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTG
GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA
ACACGTACACCAGGATGGTGCTGGAGTTCAGTGCAAGCATAGTAAGAGCTTCAAGGAAAAC
CCGAAGTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC
CGTGGGTCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGGCCAGCCAGGAGCTG
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG
TCTCTTGAGCATGGTTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC
CTAAGGGATTCTGGGTGCCACTGCTCTCTTTTCCCTCTACAGCTCCATCTTGTTTCACCCAC
CCCACATCTCACACATCCAGAAATCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC
TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP
VSLDGDNTNTSTQEVVQYNWETGDDRFSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR
GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLSPPLGLCGKNPMVI PGNADHLHRTSIHQ
PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLTTCLE
LWLGLLHGLALLHLLHGVGCHHLQHVVHODGAGVQVQA

FIGURE 191

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG
TCTTTATGTCTTTCTCCTCTTCCTATTCTGTTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA
GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG
TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT
GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG
AGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTATTCTCCTCCCAA
GTCTGTTCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCAATGGCATTACAGAGAAAG
CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT
AGCCACCTCCCTGTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTAG
GACATTCGCCCCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCT
CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT
ATGGCGATGGCCATGATGTTACAATCCCACCTGCCTGAATAATCAAGTGGGAAGGGGAAGCA
GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAA
GGGAAGCAACAGGAACCTTCTGCAACTGGTTTTTATCGGAAAGATCATCCTGCCTGCAGATGC
TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA
ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGG
AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG
AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC
CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

192/330

FIGURE 192

MWLPLGLLSLCLSPILSSPSLKSQLQQLWTLPSPLVAFRANRTTYVMDVSTNQGSME
HRNHLCFCDLYDRATSPPLKCSLL

09941-10000

GTAGCGCGTCTTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCCCTCGGGTCGTGGAGCCAGGACGACGACGTCA
CCGCC**ATGGC**AGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTTGATGCTT
GGATGTGCCCTTCCAATATACAACAAATACTGGCCCTCTTTGTTCTATTTTTTACATCCTTTCACCTATTCC
ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACTTGCCATCTTTCT
TTACAACGGGCTTTGTCGTGTACGTTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA
GCTTGTGCACCTTGTCTCACAGGAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG
CAATGACGACTTCAGCTGGCAGCAGTGG**TGA**AAAGAAATTACTGAACATTGTCAAATGGACTTCCTGTCATTT
GTTGGCCATTACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTTA
GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTACAGAGACTTGCTGAAGGATTAAGGATTCTTCT
CTTTTGGAAAAGCTTGGACTGATTTACACATTATCTATAGTATGCTTTTTGTGGTGTCTGCTGAATTTAAATAT
TTATGTGTTTTTCTGTTAGGTTGATTTTTTTTGAATCAATATGCAATGTTAAACACTTTTTTAAATGTAATCA
TTTGCATTGGTTTAGGAATTCAGAATTCGGCCGGCTCTATTACTGGTCAAGTACATCTTTCTCTTAAATATT
TAGCCTCCATTATTACAAAAAATTATAAAAAATGTTTTTCAGTCAGTCAGGATGACACTCCCAATTGTTATG
CAGACATACAGACGGTTGGCATACGTTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG
GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTG
AAAATTATTTTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCTCAATTGTTAGAAGAATTTATGTTAAACTTTA
AGGTAAGGGTGTA AAAACATTTTTTGAGATAAGGTTTTATTATTGTTTATTATTGTTAGAGTGAGTTGCAATGT
GGGAAGAAATGACATTTGAAATTTCCAGTTTTGAATCCTGTTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC
AACCTTTTCATGTTTTACCTGTTTAAATGGACATACATGGAACCACTACTGATGAGGACAGCTTGATGTTTGC
ATCATATATGCCAGAAAACCTTCCCTCTGCTTCTCTCTTTTGACTTATTTGGTATGTTGTATATATTACATAAAA
TAACTTTTCAAATATAGTTTAATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATT
CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCATGATTAACAAGTAACCTTGTTAGTCTTACAGATAATTCATGCA
TTAACGTTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAAG
TATTTTTAAGACAAGTTTCTGTATACCTCTGAACGTGTTTTGATTTTGAGTTTCATCATGATAGATCTGCTGTTT
CCTTATAAAAAGGCATTTGTGTGTGAGTTAATGCAAAGTAGCCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT
ACCTGACCAAAAAATTCCCAGTAACCAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGA
CTTTTTTTCAGGAGTGGGTTATAAAAACATTCAAGTTGGTCTGACAGTATTTTGTTAAGGATATTTGTTTGTATG
TTTATTCAGTATACCTTACATAAAAAATTATTTGCCCATCAGCCAAAACCTCAGTAATCATGACAGCTGTCTGTTGT
TTTATGAAGTTTTATTTCTCAAGAAAAATGGGAATAAAATTTGGGATTTGTTTCAGCTTTTTTACTAAAGATGCCTAA
AGCCACAGTTTTATTTGCTTAACCTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG
GCGTGTGGCTGGAGCCTCCCACTGGAGGCTGAAAGTGGCTTGTGTTATATAATGTTTCAGATTTCAAGAGGAA
GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACCTCTTGTGCTTGTGATCTACTGGACTTT
TTTTTTGCAGGAAGTGCATTCTCTGGTCTTCCCTATTTTCTGTTCTGGATGTCAGTGACGTGCAGTGCCTACTG
TTTTATCCACTTGGCCACAGACTTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT
GTGCTCTTTGACCTTGATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT
TTTCCATAGAAATGCACTGATACAACATTACCATCTCTATGAAAGAAAAACTTTTGATGATGAAACAATAA
AGATTTTAAATATCTATTTTTAAAAA

194/330

FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM
SNACKELAIFLTGTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFLVFGSND
DFSWQQW

09990411501
T05T T08550

[illegible]

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRI FQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLISLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVYAFHLCEDGA EPTSG
HLLSAVTNRSDLGPPASSATT LADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH
IEGALVIINEYGSCTCHQQPARECEV

FIGURE 197

GTGCAAGGAGCCGAGGCGAGATGGGCGTCCTGGGCCGGGTCTGCTGTGGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTGCGAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGGCGCCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTTCCGCCTAGTGCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAG
CCCCGTGCGTGTCCGCAGCATCTCGGCTCTGGGCCGGACGTTACGCGCGACGAGGACCTGG
CTGTTTTCTGGCGTCCCGCGCGGGCCGCCTACGCTTCCACGGGCCGGGCGCGCTGAGCGTG
GGCCCCGAGGACTGCGCGGACCCGTCGGGCTGCGTCTGCGGCAACGCGGAGGCGCAGCCGTG
GATCTGCGCGGCCCTGCTCCAGCCCCCT

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE
GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA
PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASR
AGRLRFHGPALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

FIGURE 199

ATCGCATCAATTGGGAGTACCATCTTCCTC**ATG**GGACCAGTGAAACAGCTGAAGCGAATGTT
TGAGCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTT
CTGCCTTTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCA
TTGACGTGGTACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTTGC
CGTGTGTCTTGCA**TAA**TTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAG
CTGGTGGACAGTTTTGTAACCTATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCT
TGCAGCAATGTGTTGCTTGTGATTCGAACATTTGAGGGTTACTTTGGAAGCAACAATACAT
TCTCGAACCTGAATGTCAGTAGCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAA
TCTTCCTCATGTACCTGTTTCCTCTCTGGATGTTGTCCCACTGAATCCCATGAATACAAAC
CTATT CAGCAACAGCAA
AAAAAAAAAAAAAAAAAAAAAAAAA

200/330

FIGURE 200

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYLSFIP
FARDAVKKCFVCLA



T63T7402660

TTGAGCGCAGGTTGAGCTCTCTGCGCGTTCCGGGGGCGTTCTCTCCAGTACCCTCCCGCGTTTACC CGCGGCGCGC
CCGAGGGAGTCTCTCCAGACCCTCCCTCCCGTTGCTCCAAACTAATACGGACTGAACGGATCGCTGCGAGGGT
GGGAGAGAAAAATTAGGGGGAGAAAGGACAGAGAGAGCAACTACCATCCATAGCCAGATAGATTATCTTACACTG
AACTGATCAAGTACTTTTGAAAATGACTTCGAAATTTATCTTGGTGTCTTTCATACTTTGCTGCACTGAGTCTTTC
AACCACCTTTTCTCTCCAACATAGACCCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT
TATATAAAGTTTCCAAGCCCCATTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT
TTTATTACAAAAACCTACCTTAACCATTATACTTTTGGTAACTGGCCTCTTTGCGAGAGAATCATGGGATTGTTGC
AAATGATATGTTTGATCCTATTCGGAACAAATCTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTT
GGGAAGAAGCGACACCAATATGGATCACAAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGA
ACAGATGTAAAAATACATAAGCGCTTTTCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG
AGTTGCCAAAAATGTTGAATGGTTTACGTCAAAAGAGCCCAATAAATCTTGGTCTTCTCTATTTGGGAAGACCCTT
ATGATATGGGCCACCAATTTGGGACCTGACAGTCCGCTCATGCGGCTGTCAATTCAGATATTGACAAGAAGTTA
GGATATCTCATACAAATGCTGAAAAAGGCAAAAGTTGTGGAACACTCTGAACCTAATCATCAAGAATGATCATGG
AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCACTATAACCCTGATTG
ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAATTTGATGAAGTCTATGAAGCACTAACTCACGCT
CATCCTAATCTTACTGTTTACAAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATAACAACAGTCGAATTCA
ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACAATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC
ACGGTTACGATAATGCGTTAGCAGATATGCATCCAAATTTTATGCCCATGGTCTGCTTCCAGAAAGAATTTT
TCAAAAGAAGCCATGAACCTCCACAGATTTGTACCCACTACTATTGCCACCTCCTCAATATCACTGCCATGCCACA
CAATGGATCATTTCTGGAATGTCCAGGATCTGCTCAATTCAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA
CTATACTCCTCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTATACCCTTATTTCATAGGGGTC
TCTCTTGGCAGCATTATAGTGATTGTATTTTTTGTAAATTTTCATTAAGCATTAAATTCACAGTCAAAATACCTGC
CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCCTAATGTTTACTTTGAAGTGATTGTCATA
TTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAATTTCTGGAAACCAAGTCCAAACATCTGC
AGAAACCATTAAGCAGTTACATATTTAGGTATACACACACACACACACACATACACACACAGGACCAA
ATACCTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGTCTCCATTGTTCACTGTAGCATAGGGATAGATAAG
ATCCTGCTTTTATTTGGACTTGGCGCAGATAATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATAT
ATTGCACTTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCATTTTATGGACAGTTTATGTCTTATAAT
TTGATTGAAAAAGACAACCTTTTTCAGCCATGTACAGAATACTTGTTCAGCAATTTGTTCAAGTGAAGGAAT
TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTTGAGAGAAGAAGAAGGTGATAAGTGTGA
AAATTAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATCCCAACAGCAGAATGCAACTGTGGGCAT
TTCTTGTCTTATTTCTTTCCAGAGAACGTGGTTTTTCATTTATTTTCCCTCAAAGAGAGTCAAATACTGACAG
ATTCTGTCTAATAATATTGTTTCTGTCAATAAATTATTGTGATTTCCTGATGAGTCATATTACTGTGATTTTCA
TAATAATGAAGACACCATGAATATACTTTTCTTCTATATAGTTTCAGCAATGGCCTGAATAGAAGCAACCAGGCA
CCATCTCAGCAATGTTTCTCTTGTGTAATTTATTTGCTCCTTTGAAAATTAATCACTATTAATTACATTA
AAATCAAAATTTGGATATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLVVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQ RAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTELIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPII AVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTTQSTILLPGSVKPAEYDQEGSYPIYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 203

GGATTTTGTGATCCGCGATTGCTCCACGGGCGGGACCTTTGTAAGTGCAGGAGGCCAG
 GACAGGCCCACCTGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGC
 AGAGAGGCCAAGCCCCCTTGCCTTGGGTACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA
 CCAGATCCAGAGGCAACAGGGAC**ATG**GCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC
 AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA
 CTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGG
 AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCCTGACGTTGCC
 CCTGCCCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAACTGTT
 CAGCTCCCACAGGTTTCAGGTCATCATCTGCTTGGTGGTCTGGATGCCCTCCTGGTGC
 TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG
 GTATTCCACTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCTTTAAATT
 ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCCGTCGTGGTGG
 TGGTCTCATTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC
 CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT
 TAAGACACGTTTCAGAACGGCAACTCTTAAGGTTAAACAGATGAATGTACAATTGGCCGCCA
 AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCCTGGACT**TGA**TGAGTTTGCTGTATC
 AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT
 CACACAGCCACCGTGAAAGTCCTGGAGTAAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAG
 CAGGCTGGCATGTTCACTGGGCTGGTGTACGACAGAGAACCCTGACAGTCACTGGCCAGTTA
 TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC
 AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEEQPPPTPV
SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSHRFQV I I ICLVVLDALLVLAELILDL
KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLLSSFTTSLRSWMPVVVVVSFILD I
VLLFQEHQFEALGILLILLRLWRVARIINGIIISVKTRSERQLRLKQMNVQLAAKIQHLEFS
CSEKPLD

FIGURE 205

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCCAGCAGGCCGCCAAC**AT**
GCTCTGTCTGTGCCTGTACGTGCCGGTTCATCGGGGAAGCCAGACCGAGTTCAGTACTTTG
 AGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCTCC
 CAGGAATTCTCCACCTACCGCCAGTGGGAAGCAGAAAATTGTACAAGCTGGAGATAAGGACCT
 TGATGGGCGAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA
 GGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG
 CAGTCCCTGCCGGGACTTGGGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAG
 CATGGATAAAAAACGGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC
 ACCCCGTGGAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCCACGATCTTTGATGTG
 GGTGAGAATCTAACGGTCCCGGATGAGTTTACAGTGGAGGAGAGGCAGACGGGGATGTGGTG
 GAGACACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGG
 ACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT
 GGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAA
 CGTCCCTCAAAATTGCCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCC
 TTGTTGGTAGTGACCAGGAGACTCTGAGGATTACGAGAGGCTTGTGGCAGGGTCTTTGGCA
 GGGGCCATCGCCAGAGCAGCATCTACCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCG
 GAAGACAGGCCAGTACTCAGGAATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGG
 TGGCGCCTTCTACAAAGGCTATGTCCCCAACATGCTGGGCATCATCCCCTATGCCGGCATC
 GACCTTGCAGTCTACGAGACGCTCAAGATGCCTGGCTGCAGCACTATGCAGTGAACAGCGC
 GGACCCCGGCGTGTGTTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGG
 CCAGCTACCCCTGCCCCAGTACAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCT
 CCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAGGGGGCCTTCGGGCT
 GTACAGGGGGCTGGCCCCAACCTCATGAAGGTTCATCCAGCTGTGAGCATCAGCTACGTGG
 TCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCCGGT**G**TCAGGGGGGAGGGCCGCCG
 GCAGTGGACTCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTTCTGTGAATG
 TGCCAACACTAAGCTGTCTCGAGCCAAGCTGTGAAAACCCTAGACGCACCCGCAGGGAGGGT
 GGGGAGAGCTGGCAGGCCCAGGGCTTGTCTGCTGACCCAGCAGACCCCTCCTGTTGGTTCC
 AGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAGCAGGCTCCGGGCTCACATGTGTAA
 GGACAGGACATTTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGCCGCTTAGT
 TCTTCCATTTACCCCTTGCAGCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGC
 ATCTCCCTGTGCCCTCTTGTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAG
 CCCACATCCCACCCCTCGTCCAATCCCATAATCCATGATGAAAGGTGAGGTCACGTGGCCT
 CCCAGGCCTGACTTCCCAACCTACAGCATATTGACGCCAACTTGGCTGTGAAGGAAGAGGAAAG
 GATCTGGCCTTGTGCTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATGCT
 TGGGAGTGCAGGGGGCTCGGGCTGCCTGGCTGGCTGCACAGAAGGCAAGTGCTGGGGCTCA
 TGGTGCTCTGAGCTGGCCTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAAACTCACTG
 TCCCCACTGTGGCATGAGGGCAGTGGAGCACCATGTTTGAGGGCGAAGGGCAGAGCGTTTGT
 GTGTTCTGGGGAGGGAAGGAAAAGGTGTTGGAGGCCTTAATTATGGAAGTGTGGGAAAAGGG
 TTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
 GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTCCAACC
 CCAGCAGGGGGCGCAGCGGGACCAGCCCCACATTCCACTTGTGTCACTGCTTGGAACTATTT
 ATTTTGTATTTATTTGAACAGAGTTATGTCTAACTATTTTTATAGATTTGTTTAATTAATA
 GCTTGTCATTTTCAAGTTCATTTTTATTTCATATTTATGTTTCATGGTTGATTGTACCTTCCC
 AAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAGGGGGCCCTTGGGCCGCTGCAGTCACAT
 CTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAAGCAGCAGCCCTG
 GCTCCTTTCTTTTGGCAGGTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTTCAAGGTTTGA
 CTGGGGGGCGTGGAGAGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTT
 CTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCT
 CACTGTGAATTTTGTGGTGGGCGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGCTCCCTCC
 CAGCCTTCTGCTGCCCTTGTCTTAACAATGCCGGCCAACCTGGCGACCTCACGGTGCACCTCC
 ATTCCACCAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATT
 GTTATATATGAACATATAACTGGAGTCGTCAAAAAGCAAATTAAGAAAGAATTGGACGTTAG
 AAGTTGTCAATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAA
 AA

FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD
LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILK
SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW
WRHLVAGGGAGAVSRTCTAPLDRLKVLMMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI
NVLKIAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMAL
RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNS
ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQAASIEGAPEVTMSSSLFKHILRTEGAFG
LYRGLAPNFMKVIPAVSISYVVYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation site.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 208

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI
KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRТАVFADQVIVGNASRLKNVQLTD
AGTYKCYIIITSKGKGNANLEYKTGAFSMPEVNVNDYNASSETLRCEAPRWFPQPTVWASQVD
QGАНFSEVSNTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRR
SHLQLLNSKASLCVSSFFAISWALLPLSPYMLK

FIGURE 209

GAATTTGTAGAAGACAGCGGCGTTGCCATGGCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT
GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC
AGCGGGTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG
AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC
CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACCTCTCTGGCTATCATCTTCA
CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC
GAGTGCGGGACGCAGCTCTGTGGATCTCGACATACTGTGTTAGTTCCTTCCCAGAACCCAT
CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCCTCTGCAGCTGTTTT
GCTTCCTTGTGGCCATCAGAGTTCCCTTCCCTTGGACAGTCTGGAGAAAGACAGAGGCTGGG
GTTTGGGATTGAAGACCAGACCCCATCTGAGCCCTTCCTCCAGCCCTGTACCAGCTCCTACT
GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT
GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC
AACAGTCTACCCTTTGAGTGGGCCGAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCT
GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT
CCTCCTCTCTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG
AGTTTTGTAACTTCAAGTGCTGTTCACTGCGGGGATTTAGCACAGGAGACTCTACGCTCA
CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCTGCTAACATCTCAGGCTCCCAGCCCA
GCCACCATTAAGTGCTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT
CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG
GGAAGTCACTGAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA
GTGGCAGATGCAGGAAATGAGCTGTCTATTAGCCTTGCTGCCCCACCCATGAGGTAGGCAG
AAATCCTCACTGCCAGCCCCCTCTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCCACCTGAC
TCCAGCACACCTGGCGAGTAGTAGCTGTCAATAAATCTATGTAAACAGACAAAAAAAAAAAAA
AAA

210/330

FIGURE 210

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLQEMKTLFLNTEYLMPFLLNQCSSLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGSRHTCVSSFPEPISPEWVRTRPFPILPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 211

3.

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLET PVR
LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE
CPACYESNGTSCRGKPWKYEEEEQCVFLVAELKNDIESKSLVLKGC SNVSNATCQFLSGENK
TLGGVIFRKFECANVNSLTPTSAPT TSHNVGSKASLYLLALASLLLRLGLLP

FIGURE 213

GGCCTTCGGTTCAAACGACCCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCA
GGGCTTGCCTCACTGGCCACCCCTCCCAACCCCAAGAGCCAGCCCC**ATG**GTCCCCGCCGCCG
GCGCGCTGCTGTGGGTCTGCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCCAAGGCCTG
ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTTGGGGGCCCCATGACCCGCAG
CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA
ATGATGCCATGGCCGACGCCGACCGCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC
ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA
GGGGGTGTGATTAATGCCGGAAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCCA
ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG
ACTTCAAGCCTGCCGCGCTCCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC
CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCAACCCTACCCACAG
CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCCGTGGCACTGCCAC
TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCGGGGCGCCT
TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGACCTATCAACAATGTC
CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT
GCCTCTCAGAGCACCACCAGTACCAGGACCACCCTACCCCTTCCCCACCATCCACCTCAG
AAGCAGTCCCAGCCTGCCACCCGCCAGCCCCTGCCAGCCCTGGCTTTTTTGGAACGGGTCA
GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATA
GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT
CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAAACCTGGAAAA
CACAA

MVPAAGALLWVLLLNLGPRAGAQLTQTPTMQRVSLRFGGPMTRSYRSTARTGLPRKTRI
ILEDENDAMADADRLAGPAAAEALLAATVSTGFSSAINEEDGSSEEGVVINAGKDSTSREL
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP
SPSPTAMPSPEDLRLVLPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC
TYQQCPNRLREECPLDTSLCDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA
FWKRVIRIGLEDIWNLSVVFTEMQPIDRNQR

FIGURE 215

CCGGGGTCGACCCACGCGTCCGGGGAGAAAG**GATG**CGCGGCCTGGCGGCGCGGTTGGTCCTGCTAGCTGGGGCA
 GCGGCGCTGGCGAGCGGCTCCCAGGGCGACCGTGAGCCGGTGTACCGCGACTGCGTACTGCAGTGCGAAGAGCA
 GAACTGCTCTGGGGGCGCTCTGAATCACTTCCGCTCCCGCCAGCCAATCTACATGAGTCTAGCAGGCTGGACCT
 GTCGGGACGACTGTAAGTATGAGTGTATGTGGGTACCGTTGGGCTCTACCTCCAGGAAGGTCACAAAGTGCCT
 CAGTTCCATGGCAAGTGGCCCTTCTCCCGTTCTCTTTCAAGAGCCGGCATCGGCCGTGGCCTCGTTTCT
 CAATGGCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCATGTACCACA
 CCTGTGTGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACCAGGGACACTGAC
 CTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCTACCTCAATCTACCTGTGCTGCGTCAGGAC
 CGTGGGGCTGCAGCACCCAGCTGTGGTCAAGTGCCTTCCGGGCTCTCCTGCTGCTCATGCTGACCGTGCAGTCT
 CCTACCTGAGCCTCATCCGCTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTG
 GTGTGGTGGCTGGCCTGGTGCCTGTGGAACCAGCGCGGCTGCCTCACGTGCGCAAGTGCCTGGTGGTGGTCTT
 GCTGCTGCAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCACCGCTCTTCTGGGTCCTGGATGCCCATGCCA
 TCTGGCACATCAGCACCATCCCTGTCCACGTCTCTTTTTAGCTTTCTGGAAGATGACAGCCTGTACCTGCTG
 AAGGAATCAGAGGACAAGTTCAAGCTGGAC**TGA**AGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCC
 GCCCTGCTGGCCTCCCTTCTCCCTCAACCTTGAGATGATTTTCTCTTTTCAACTTCTTGAACCTGGACATGA
 AGGATGTGGGGCCAGAATCATGTGGCCAGCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGG
 AAGGCCTCCAGCATCTGGGACTCGAGAGTGGGCGAGCCCTCTACCTCCTGGAGCTGAACCTGGGGTGGAACTGA
 GTGTGTTCTTAGCTCTACCGGGAGGACAGCTGCCTGTTTCTCCCCACCAGCCTCCTCCCCACATCCCCAGCTG
 CCTGGCTGGGTCTGAAGCCCTCTGTCTACCTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCC
 CTTCTGTTACCACCCCCACCCCTCCTCCAGGACACCCTAGGTGGTGTGCTGGATGCTTGTTCTTTGGCCAGCCAA
 GGTTACGGCGATTCTCCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTTGGGAAGGAGTTTCACCCTGACC
 GTTGGCCTAGCCAGTTCCAGAGGCTCACCATACTCCCTTTCAGGGCCAGGGCTCCAGCAAGCCAGGGCA
 AGGATCCTGTGCTGCTGTCTGTTGAGAGCCTGCCACCGTGTGTGCGGAGTGTGGGCCAGGCTGAGTGCATAGG
 TGACAGGGCCGTGAGCATGGGCTGGGTGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTTGT
 CGGGGAAGAGGTGTGGCTTCAAAGTGTGTGTGTGCAGGGGTGGGTGTGTTAGCGTGGGTAGGGGAACGTGTG
 TGCGCGTGTGTTGGGCATGTGAGATGAGTACTGCCGTGAATGTGTCCAAGTTGAGAGGTGTGGAGCAGGAT
 GAGGGAATCCTGTCAACATCAATAACTTGTGGAGCGCCAGCTCTGCCCAAGACGCCACCTGGGCGGACAGC
 CAGGAGCTCTCCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGGTGGCCCTTTGCCCGCCTCCTGCAAC
 CTCACAGGGTCCCCACACAACAGTGCCTCCAGAAGCAGCCCTCGGAGGCAGAGGAAGGAAAATGGGGATGGC
 TGGGGCTCTCTCCATCCTCTTTTCTCCTTGCCCTTCGCATGGCTGGCCTTCCCCTCCAAAACCTCCATTCCCCT
 GCTGCCAGCCCTTTGCCATAGCCTGATTTTGGGGAGGAGGAAGGGGCGATTGAGGGAGAAGGGGAGAAAGCT
 TATGGCTGGGTCTGGTTTCTTCCCTTCCCAGAGGTCTTACTGTTCCAGGTGGCCCCAGGGCAGGCAGGGGCC
 ACATAATGCCTGTGCCCTGGTAAAGGTGACCCCTGCCATTTACCAGCAGCCCTGGCATGTTCCCTGCCCCACAGG
 AATAGAATGGAGGGAGCTCCAGAACTTTCCATCCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTG
 CTCTGCCCTGACCCCTTGTCCCTCTTTGAGGGAGGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTG
 GCCTGCGCTAGCTTCTTTTGATACTGAAAACCTTTAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAA
 TTCCAAGCCTCAAAAAAAAAAAAAAAAAA

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGGLYLQEGHKVPQFHGKWPFSSRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAWFWSTVFHTRDTDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYSLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLQGLSLELLDFPPLFWVLDAAHAIWHISTIPVHVLFFSFLEDDSLYLL
KESEDKFKLD

Important features:**Signal peptide:**

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

FIGURE 217

[illegible]

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPESSTKETERKETKAEHEEL
 DAEVLEVHFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGGKRLDINTN
 TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
 RLINKFNSSSSSLEEKIAALFDLEYVHQM DNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
 VLGAAFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
 GGLQVLR TLVQEKGTEVLAVRVV TLLYDLVTEKMF AEEEEAE LTQEMSPEKLQQYRQVHLLPG
 LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
 LELQDGEDEGYFQELLGSVNSLLKELR

Important features:**Signal peptide:**

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

FIGURE 219

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCTTCCC
 CTTCCCCGGGGTCTGGGGGTGACATTGCACCGCGCCCCCTCGTGGGGTCGCGTTGCCACCCCA
 CGCGGACTCCCCAGCTGGCGCGCCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCTTCC
 TTCCACCTGACCAGCC**ATG**GGGGGCTGCGGTGTTTTTTCGGCTGCACTTTTCGTCGCGTTCGGC
 CCGGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGT
 CGCAGGGGCATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG
 TCCATGTGACCGACCGGTGAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCT
 GTCTCTGTCCTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGA
 TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT
 ATGTTTCTGGTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCT
 GATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCTATTACTTCCTGACTTC
 AGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATG
 CCTGTGAGAGGAGACGGTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCG
 GGACTGACATTCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGT
 TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTCAGCGCAGCC
 TCTTGTGTAAGGACT**TGAC**TACCTGGACTGATCGCCTGACAGATCCCACCTGCCTGTCCACTG
 CCCATGACTGAGCCCAGCCCCAGCCGGGTCCATTGCCACATTCTCTGTCTCCTTCTCGTC
 GGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCTTTTGTGACCGTTAGTCTCTAAGCTT
 TACCAGGAGCAGCCTGGGTTGAGCCAGTCAGTGACTGGTGGGTTTGAATCTGCACTTATCCC
 CACCACCTGGGGACCCCTTGTGTTGTGTCCAGGACTCCCCCTGTGTGAGTGCTCTGCTCTCAC
 CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGGCAGGAGGACAGTCGGGTGAT
 GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC
 CTGGGGTGCCCTCCTGATGTCCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG
 GTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAA
 GGATAGATGAGCTCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC
 AGGCCTGAGGGGGAACCATTTTTTGGTGTGATAAATACCCTAAACTGCCTTTTTTTCTTTTTT
 GAGGTGGGGGGAGGGAGGAGGTATATTGGAACCTCTCTAACCTCCTTGGGCTATATTTCTC
 TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTCCGTCCCTTCTCCTTGGTCCCAGACCTT
 GGGGGAAGGAAGGAAGTGATGTTTGGGAACCTGGCATTACTGGAACATAATGGTTTTAACCT
 CCTTAACCACCAGCATCCCTCCTCTCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGC
 CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGG
 GAGATTTTTTTGTAGTTTTTAATTGGGGTGTGGGAGGGGCGGGGAGGTTTTCTATAAACTGT
 ATCATTTTTCTGCTGAGGGTGGAGTGTCCATCCTTTTAATCAAGGTGATTGTGATTTTGA
 AATAAAAAAGAATTTGTAAAAA
 AA
 AA

FIGURE 220

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIIILVAGAFFWLVSLLLASVWVFILVHVTDR
SDARLQYGLLI FGAAVSVLLQEVRFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS
FGIISGVFSVINILADALGPGVVG IHGDSPIYFLTSAFLTAAIILLHTFWGVVFFDACERRR
YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLCKD

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 221

AAGCTGGTTTAAGGAAGCAGAGGAGGGTTAGATTCGTTGAGTGAGGACGGAAGATCAACCCA
TTTCCATTCCGCCAGATGGCCTATGTTTCTGGTCTCTCCCTTCGGNATCATCAGTGGTGTNT
TNTCTGTTATCAATATTTTGGCTGATGCANTTGGGCCAGGTGTGGTTGGGATCCATGGAGAC
TCACCCTATTANTTCCTGAN TTCAGCCTTTNTGACAGCAGCCATTATCCTGCTC

1099391403660

FIGURE 222

GACCGACCGTTCAGATGCCCGGTTCCAGTACGGCTTCCTGATTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTTTTNTCTGTTATCAATATTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCTATTAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCCTTTCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTCCTGTCCTGGTCAGGCCCCACCCCCCTTCCCACNTG
ACCAGCCATGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC
CGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCNCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCCC
CAGNTGGCGCGCCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTGCGGTTCTGGGCCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA
CCGACCGGTGAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

FIGURE 225

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCCGGTGCCCAGAGCCCAGGAGGAGGCAG
 TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCC
 TACCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGG
 TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCCTCCTTGCCCTGTCTGGAGGCTGCT
 AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGT
 CCTTGTGGTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCCATGGGCTACAGCAAGACCCC
 CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTT
 CTCGCCAACAAATGATGTTTCCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA
 GGACCTGGGAGCTGGGGCCGGGGGAAGACGCCCGGTGCGATGACAGCAGCAGCCGCATCATCA
 ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC
 CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCGAG
 GAAGAAAGTTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC
 AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCTGGCTACTCCACCCTGGCCACTCT
 AACGACCTCATGCTCATCAAACCTGAACAGAAGAATTTCGTCCCACTAAAGATGTCAGACCCAT
 CAACGTCTCCTCTCATTTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAA
 CCAAGAGCCCCCAAGTGCACCTTCCCTAAGGTCCCTCCAGTGCTTGAATATCAGCGTGCTAAGT
 CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA
 CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC
 TGCAGGGACTCGTGTCTGGGGAGATTACCCTTGTGCCCGGCCCAACAGACCGGGTGTCTAC
 ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAT
 CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTTCAG
 ACCCTCATTCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCCTGACCCCATGTCT
 CCTGGACTCAGGGTCTGCTTCCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG
 GAACAATTTCCAAAACCTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCAT
 CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA
 CTGAGAAGTGGAAAAAAAAAA

FIGURE 226

MATARPPMMWVLCALITALLLGVTEHVLANNVSCDHPSNTVPSGSNQDLGAGAGEDARSDD
SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS
PVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCP SAGTKCL
VSGWGTTKSPQVHFPKVLQCLNISVLSQKRCE DAYPRQIDDTMFCAGDKAGRDSCQGDSGGP
VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

protein

ATGGTCAACGACCCGGTGGGAAGACCATGGGCGGCGCTGCCCAACTTGAGGACCGGGCCGCGCA
CAAGCCGCGAGCGGCCGAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC
TGCTGGCTGTAGCTGTGTACCGGTGCCGTGCTCTTCTGAACCACGCCACGCGCCGGGCACG
CGCCCCCACCTGTCTGTACGACTGGGGCTGCCAGCGCCAACAGCGCCCTGGTCACTGTGGA
AAGGGCGGACAGCTCGCACCTCAGCATCTCTATTGACCCGCTGCCCGACCTCACCAGCA
GCTTCGCACGCCTGGAGAGCGCCCAGGCCTCGGTGCTGCAGGCGCTGACAGAGCACCAGGCC
CAGCCACGGCTGGTGGGCGACCAGGAGCAGGAGCTGCTGGACACGCTGGCCGACCAGCTGCC
CCGGCTGCTGGCCCGAGCCTCAGAGCTGCAGACGGAGTGCATGGGGCTGCGGAAGGGGCGATG
GCACGCTGGGCGAGGCCCTCAGCGCCCTGCAGAGTGAGCAGGGCGCGCTCATCCAGCTTCTC
TCTGAGAGCCAGGGCCACATGGCTCAGCTGGTGAACCTCCGTACGACATCTGGATGCCCT
GCAGAGGGACCGGGGGCTGGGCGGGCCCCGCAACAAGGCCGACCTTCAGAGAGCGCTGCC
GGGGAACCCGGCCCCGGGGCTGTGCCACTGGCTCCCGGCCCGAGACTGTCTGGACGTCCTC
CTAAGCGGACAGCAGGACGATGGCGTCTACTCTGTCTTTCCACCCACTACCCGGCCGGCTT
CCAGGTGTACTGTGACATGCGCACGACGGCGGCGGCTGGACGTTTCAGCGCCGGGAGG
ACGGCTCCGTGAACTTCTTCGGGGCTGGGACGCTACCGAGACGGCTTTGGCAGGCTCAC
GGGAGCACTGGCTAGGGCTCAAGAGGATCCACGCCCTGACCACACAGGCTGCCTACGAGCT
GCACGTGGACCTGGAGGACTTTGAGAATGGCACGGCCTATGCCCGCTACGGGAGCTTCGGCG
TGGGCTTGTTCTCCGTGGACCCTGAGGAAGACGGGTACCCGCTACCGTGGCTGACTATTCC
GGCACTGACGGCGACTCCCTCCTGAAGCACAGCGGCATGAGGTTCACCAAGGACCGTGA
CAGCGACCATTTCAGAGAACAACCTGTGCCCTTCTACCGCGGTGCCTGGTGGTACCGCACT
GCCACACGTCCAACCTCAATGGGCAGTACCTGCGCGGTGCGCACGCCTCCTATGCCGACGGC
GTGGAGTGGTCCCTCCTGGACCGGCTGGCAGTACTCACTCAAGTTCTCTGAGATGAAGATCCG
GCCGTCCGGGAGGACCGCTAGACTGGTGCACCTTGTCCTTGGCCCTGCTGGTCCCTGTGCG
CCCCCCCCGACCCCTACTCTTCTGTAATGTTCTTCCACCCACTGTGCCTGGCTGGCGGAC
CCACTCTCCAGTAGGGAGGGGGCGGGCCATCCCTGACACGAAGCTCCCTGGGCGGGTGAAGT
CACACATCGCCTTCTCGCCGTCCCCACCCCTCCATTTGGCAGCTCACTGATCTCTTGGCTC
TGCTGATGGGGGCTGGCAAACCTTGACGACCCCAACTCCTGCCTGCCCCACTGTGACTCCGG
TGCTGTTTGGCGTCCCCTGGCCAGGATGGTGGAGTCTGCCCCAGGCACCCCTCTGCCCTGCC
GGCCAAATACCCGGCATTATGGGGACAGAGGAGCAGGGGGCAGACAGCACCCTGGAGTCTC
CTAGCAGATCGTGGGGAATGTGAGGTCTCTCTGAGGTGAGGTCTGAGGCCAGTATCCTCCAG
CCCTCCCAATGCCAACCCCCACCCCGTTTCCCTGGTGCCAGAGAACCACCTCTCCCCAA
GGGCTTCAGCCTGGCTGTGGGCTGGGTGGCCCCATCCTACAGGCCCTGAGGTGAGGATGGG
GAGCTGCTGCCTTTGGGGACCCACGCTCCAAGGCTGAGACGATTCCCTGGAGGCCACCCAC
CCTGTGCCCGGACAGGCTGGGGTCTGCACTCCTCTTACCTGCTGTGCCACCTGCTCTCTG
TCTCAAATGAGGCCCAACCCATCCCCACCCAGCTCCCGGCCGTCTCTACCTGGGGCAGC
CGGGGCTGCCATCCCATTCTCCTGCCTCTGGAAGGTGGGTGGGGCCCTGCACCGTGGGGCT
GGACTGCGCTAATGGGAAGCTCTTGGTTTTCTGGGCTGGGGCTAGGCAGGGCTGGGATGAG
GCTTGTAACAACCCCAACCAATTCCAGGGACTCCAGGTCCTGAGGCCTCCAGGAGG
GCCTTGGGGGTGATGACCCCTTCCCTGAGGTGGCTGTCTCCATGAGGAGGCCAACCTTGCC
ATTGACCGTGGCCACCTGGACCCAGGCCAGGCCCGGCCGCGAGTGGTCAAGGGACAGGGA
CCACCTCACCGGGGCAATGGGGTCGGGGGGACTGGGGCACCAGACCAGGCACCACCTGGACA
CTTTCTTGTTGAATCTCCCAACACCCAGCAGTGTATCCCACTCCTTGTGTGCACACA
TGCAGAGGTGAGACCCGACGGCTCCAGGACAGCAGCCACACAGGGCAGGGCTGGAGCCGG
TCCTCAGCTGTCTGCTCAGCAGCCCTGGACCCGCGTGCCTACGTACGAGCCAGATGCAGGG
CGGCTTTTCAAGGCCTCCTGATGGGGGCCTCCGAAAGGGCTGGAGTCAGCCTTGGGGAGCT
GCCTAGCAGCCTCTCCTCGGGCAGGAGGGGAGGTGGCTTCCCTCCAAAGGACACCCGATGGCA
GTGTCCTAGGGGCTGTGGGGTTCCGTTCTCCCTTCCCCTCCCACTGAAGTTTGTGCTTAAAA
AACAAATAAATTTGACTTGGCACCACCTGGGGGTTGGTGGGAGAGGCCGTGTGACCTGGCTCTC
TGTCCAGTGCCACCAGGTTCATCCACATGCGCAG

FIGURE 228

MVNDRWKTMGGAAQLEDPRDPKQRPSCGYVLC TVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELDLADQLPRLARASELQTECMGLRKHGHTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQRDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRRREDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLED FENG TAYARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASYADG
VEWSSWTGWQYSLKFSEM KIRPVREDR

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT
 TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCT
 CACCGCTGTAGGAATCCAG**ATG**CAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG
 ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACCTCGGCATCCAGAGCCC
 CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC
 TTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTACTACC
 AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCC
 CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC
 TGAAAAACTCTGTCTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG
 AACAAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG
 GACTGTAAATATTTCTGCCTTAGTGAAAACCTCTACCATGCTGAAGATAAACAAACAAGAAGA
 CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT
 TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACTTCTGAACTG
 TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG
 GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA
 TGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGAC**TGA**ATTCGCC
 CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT
 TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG
 TTTCTGTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA
 AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG
 GCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTCAT
 GTCTTCCTTACACTTGGTGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC
 ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC
 AGCAAATACACAAGGAATTCTTTTTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCAT
 CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG
 AATCTCAAATCTCAATGCCTTATAAGCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTG
 TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCTCTCCGTTTCATATCAG
 AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA
 TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG
 GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

FIGURE 230

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSSTWRPVALTLLTLCVLVLL
IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE
LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS
QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKD
CKELKRCVCERRAGMVKPESLHVPPETLGEGD

099041-1501
"TGT" 103550

FIGURE 231

AATTTTCACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGNTGG
ATGATGATGGGACACCACCATGAGCCTGCATTNTCAAGCTTTTGCCACAATTCGGCATCCAG
AGCCCCGGCGCACAGAGCACAGGGNTCCTTTTTCAACGTGGCGACCAGTGGCCCTGACCCTG
CTGACTTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTA
CTACCAGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATA
CGTCCCAAGAGTTGCAATTTNTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTNTGCAGCAT
GTGGCTGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGGAACTTTGAAGGAGGGCAA
AGTNTCCTCATNTACTATACACACACCACTTCCC

FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGGCTC
 CGGGGATTCGGCTCGGGCCGCTGGCTCTGCTCTGCGGGGAGGGAGCGGGCCCCGCCGCGGGG
 CCCGAGCCCTCCGGATCCGCCCCCTCCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT
 CTGGGGGTTCGCCGGGGCCGGGGACCCGCGGTCCGGGCGCC**ATG**CGGGCATCGCTGCTGCTG
 TCGGTGCTGCGGCCCGCAGGGCCCGTGGCCGTGGGCATCTCCCTGGGCTTCACCCTGAGCCT
 GCTCAGCGTCACCTGGGTGGAGGAGCCGTGCGGGCCAGGCCCGCCCCAACCTGGAGACTCTG
 AGCTGCCGCCGCGCGGCAACACCAACGCGGCGCGCCGGCCCAACTCGGTGCAGCCCGGAGCG
 GAGCGCGAGAAGCCCCGGGGCCGGCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC
 CTACCACCCTGCACAGCCCGGCCAGGCCGCCAAAAAGGCCGTACAGGACCCGCTACATCAGCA
 CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC
 ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC
 GGGCGCACGGGGCCGCCGGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC
 CCATTGGACACCTGCACCTGGCGCTGCGCCACCTGCTGGAGCAGCACGGCGACGACTTTGAC
 TGGTTCTTCCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG
 CCACCTCAGCCTGGCCTCCGCCGCCACCTGTACCTGGGCCGGCCCCAGGACTTCATCGGCG
 GAGAGCCCAACCCCGGCCGCTACTGCCACGGACGCTTTGGGGTGCTGCTGTCGCGCATGCTG
 CTGCAACAACCTGCGCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCGCCCTGA
 CGAGTGGCTGGGTGCTGCAATTCTCGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG
 GGGTGCACTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGACAGGAGGGGGACCCTCAT
 TTCCGAAGTGCCCTGACAGCCACCCCTGTGCGTGACCCTGTGCACATGTACCAGCTGCACAA
 AGCTTTTCGCCCCGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA
 TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT
 CCAGCACCATCCCGCCCCGGCCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA
 GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCACTGCGTGGGGCTGACCGGGCTG
 ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG
 CGGCTCCAGAAGCAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCCGGGGTATGGA
 ATACACGCTGGACTTGACGCTGGAGGCACTGACCCCCCAGGGAGGCCGCGGCCCTCACTC
 GCCGAGTGACGCTGCTCCGGCCGCTGAGCCGCGTGAGATCTTGCCCTGTGCCCTATGTCACT
 GAGGCCTCACGTCTCACTGTGCTGCTGCCTCTAGCTGCGGCTGAGCGTGACCTGGCCCCCTGG
 CTTCTTGAGGGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC
 TGCTGCTACTGTATGAGCCGCGCCAGGCCAGCGCGTGCCCATGCAGATGTCTTCGCACCT
 GTCAAGGCCACGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGGTGCCATGGCTCAG
 TGTGCAGACAGCCGCACCCCTCACCCTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC
 TGGACACACTGTTCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC
 TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA
 CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCAGAGCTGGGCCGTGACACTGGCCGCT
 TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACCTCCGACTACGTGGCAGCCCGTGGG
 CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGAGCTGCTGGAGAGCCTGGATGTGTACGAGCT
 GTTCCTCCACTTCTCCAGTCTGCATGTGCTGCGGGCGGTGGAGCCGGCGCTGCTGCAGCGCT
 ACCGGGCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC
 GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG
 CAACAGCACCT**TGA**CCCCACCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCACCCCACTT
 CTCCCCAAAACCAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC
 AAGCTGGCCCCACTGGTCCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG
 GGACGTGCCCCCCAGAGCCACCACTTCTCATCCAAACCCAGTTTCCCTGCCCCCTGACGCT
 GCTGATTTCGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC
 CTCTGGGCCCTGGGGGCTGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGGAGCTGAGGAGGGG
 GCATCTCCCAACTTCTCCCTTTTGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA
 AGTGTGGAAAAA

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELP PRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQA AKKAVRTRYISTELGIRQRLLVAVL
TSQTTLP TLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERP IGHHLHLALRHLL E
QHGD DFDWFFLV PDDTTYEAHGLARLTGHL SLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGR CILDATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHF RSALT AHPVRDPVHMYQLHKAFARAE LERTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADV LGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALT PQGGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAF FPMHFQAFH PGVAPPQGP GPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLES LDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQE QGNST

FIGURE 234

GCTCTGGCCGGCCCCGGCGATTGGTCACCGCCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT
 TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCAGTGGAGAAAAGGAGT
 AGCTATTAGCCAATTTCGGCAGGGCCCCGCTTTTTAGAAAGCTTGATTTCCCTTTGAAGATGAAAG
 ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA
 CTGTATCCACCCAAATGTCACCGATTTCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA
 GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACCTGGAGGCAAAG
 AGGGTTGCTCAACGCCCCGCCTCATTTGAAAACCAAATCAGATCTGGGACCTATATAGCGTG
 GCGGAGGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT
 TTCCCCGCCCCTGAGACCCTGCAGCACCATCTGTC**ATG**GCGGCTGGGCTGTTTGGTTTGAGC
 GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCGCGTCCGCTGGGA
 ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC
 CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACCTTGATGAGAAGAACCCA
 GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT
 CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA
 GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC
 CTTCCCATCATGGAATCCAACCTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAG**TG**
ACCAGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCTGCCTGCCATTCTGAC
 CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

FIGURE 236

GGCGGCTGGGCTGTTTGGTTTGAGCGCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGC
TCCCGGCCGCCCCGCGTCCGCTGGGAATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCT
GTGGCGGGAAAGCGGCCCCCAGAACCGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGA
CGAAACTTGTATGAGAAGAACCCAGACTCCCATGGTTATGACAAGGACCCCGTTTTGGACG
TCTGGAACATGCGACTTGTCTTCTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACC
TTTGTGGCCTATCTGCCTGACTACAGGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCT
TGTGAAATACCGAGAGGCCAATGGCCTTCCCATCATGGAATCCAAGTCTTCGACCCCAGCA
AGATCCAG

FIGURE 237

GCGGCGGCT**ATG**CCGCTTGCTCTGCTCGTCCTGTTGCTCCTGGGGCCCCGGCGGCTGGTGCCT
 TGCAGAACCCCCACGCGACAGCCTGCGGGAGGAACTTGTATCACCCCGCTGCCTTCCGGGG
 ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTTCGGAGCTTCAGCGGGAAGGAGTG
 TCCCATTAACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA
 GCTGCACCTGTCAATCACACAAGGCTTTTGGAGGACCCGATACTGGGGGCCACCCTTCCTGC
 AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAA
 TCTTGGAAGGAGCTCAGTAATGTCTCTCAGGGATCTTCTGCGCCTCTCTCAACTTCATCGA
 CTCCACCAACACAGTCACTCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG
 ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC
 CCCTGGAAGAAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA
 TCGCTTGTTCCACACCAGCTACCACTCCAGGCAGTGCATATCCGCCCTGTTTGAGAAATG
 CACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCAGTTGTATTTGATGCCTTC
 ATCACGGGGCAGGGAAAGAAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCCTCACGGA
 GCCCTGCCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACTACAACCAGGACA
 ACGAGACATTAGAGGTGCACCCACCCCCGACCACTACATATCAGGACGTCATCCTAGGCACT
 CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACCTCTCGAAACCT
 CAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCCCCAGTGCCCTTCCTGC
 ATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTAC
 AACACCCACCCATACCGGGCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCG
 GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAACAACCAAGTTACATCC
 ACTACCAGCCTGCCCAGGACCGGCTGCAACCCACCTCCTGGAGATGCTGATTACAGCTGCCG
 GCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA
 CACGCCAGATCCTAACCATGGCTTCTATGTACGCCATCTGTCCTCAGCGCCCTTGTGCCCA
 GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCTCTTCAACAGCCTGTTCCCA
 GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC
 GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT
 GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT
 GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCACTCT**TGATT**
 CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT
 TCTGCCACTTGCTCTCCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTCAGGGC
 CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTGGAATTTGAATTAA
 CTTAGAAATTCATTTCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA
 AGTGGTCGGTGGCTGCTGTATTGGACAGCACAGAAAAAGATTTCCATCACCAAGAAAGGTC
 GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGTGTAGTGA
 TGGAGTTTACTGTTTGTGGAATAAAAACGGCTGTTTCCGTGGAAAAAAAAAAAA

FIGURE 238

MPLALLVLLLLGPGGWCLAEPPrDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTYWGPPFLQAPSGAELWVWFQDTVTDVDSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASFKPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSSISWELRQTLVVFDAFITG
QGKKDWSLFRMFSTRLTPEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLOPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLLVNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

FIGURE 239

CAACATGGGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG
CTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC
TGCTTCAAGTCCGATCCTCCCCAGTGTACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCTGTGAAGGAACTGGAAGAAGGAGGAA
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC
TCCTCCTCTACCAGGTGTCCTCAGAATGATGCTGGGTCCTTTCTACCTCTGGGGGTCACTC
TCACTTGGCACCTGCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAGAGGGAAGAGTCACAAAAAG
TCCAGACCCCAGGGACGGTACTTTCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT
GGACCCCTCATGAATGAAACCAGTGCCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAACTGGTCCTCACCAGATG
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAACTGTAAGAAATAAATAT
TTGCTGTTTATAATCCAA

240/330

FIGURE 240

MGSSSFLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC
YLHCGFKCVIPVKELEEKGKDEEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

Downloaded from www.jstor.org

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACCTTCCTGGCCAGG
AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC
TCTAGAACCCGACCCACCACCA**ATG**AGGTCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGG
CGTCCAGTGGTCTTTGCTTCTGGCTGTCCTGGTCTTCTTTCTCTTCGCCCTTGCCCTCTTTTA
TTAAGGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCT
CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA
TGCAGAGCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCAGCCCAAGGCCACACCA
CCGGAGACAGAGGAAAGGAGGCCAACCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCAC
ACAGCACAGAGGGCAGCATGGAAGAGCCAGAAAAAGAGAAAACCATGGTGAACACACTGTC
ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC
AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG
GTGTCAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAGTCA
GCACAGAATGCTGGCTCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA
CAGCAGTCATCCCACCTAAGGAGAAGAAACCTCAGGCCACCCACCCCCCTGCCCTTTCCAG
AGCCCCACGACGCAGAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG
GGATTTTGGAGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG
TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAACTCTTTCTGCCCAACCTCACTCTC
TTCTGGA

FIGURE 242

MRSCLWRCRHL SQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
 KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
 KSPEKEKTMVNTLS PRGQDAGMASGRTEAQSWKSQDTKTTQGNNGQTRKLTASRTVSEKHQG
 KAATTAKTLIPKSQHRMLAPTGA VSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN
 QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
 NQSEWDRLEHFAPPF GFMELNYSLVQKVVTFRFPVPQQQLLLASLPAGSLRCITCAVVGNGG
 ILNNSHMQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGF TAFSLTQSL LILGNRGFKNVP
 LGKDVRYLHFLEGT RDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
 RYMKNRFLRSKTL DGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSW
 KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

FIGURE 243

CGATGCGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTGCTGGGGCCTTCGCCG
GAGCAGCGAGTGGAAATTGTTCCCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAAACACCT
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA
CCAGAGACTGCACCATTCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT
GTCAGTCCAGTGTGCATGGAGGATAAGTGAGCAGACCGTACAGGAGCAGCACACCAGGAGCC
ATGAGAAGTGCCTTGGAACCAACAGGGAAACAGAACTATCTTTATACACATCCCCTCATGG
ACAAGAGATTTATTTTTGCAGACAGACTCTTCCATAAGTCCTTTGAGTTTTGTATGTTGTTG
ACAGTTTGCAGATATATATTCGATAAATCAGTGTACTTGACAGTGTTATCTGTCACCTATTT

244/330

FIGURE 244

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFRLYHNT
RDCTIPAYYKRCARLLTRLAVSPVCMEDK

EST-103650

FIGURE 245

GGGCTGGGCCCCGCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG
CCCGACCCCGGCGCGCCAGCCCCACCATGCCACCCGCGGGGCTCCGCCGGGCCGCGCCG
CTCACCGCAATCGCTCTGTTGGTGCTGGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTCCCTGTTGCTACCTGT
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTTGAAGGCCAGGAGATTCCAATGACAGGCATC
CCAGTGACGCCAGTATACCCATACCCCCAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC
CAGTCTACAACCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC
TGAGGAACCAGCCATGTCTCTGCTGCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA
GCCAAGCCCTGGGCCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA
ACTATGAGGGGTGGGGGGAGGGCTTGGAATTATGGGCTATTTTTTACTGGGGGCAAGGGAGG
GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT
TCCGTGAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG
CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRAAPLTAIALLVLGAPLVLAGEDCLWYLDRNGSWHPGFNCEFFTCCGTCYHRYC
CRDLTLLITERQQKHCLA FSPKTIAGIASAVILFVAVVATTICCF LSCCYLYRRRQQLQSP
FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYP PPSGPAPQYPLYPAGPPVYNPAAPPP
YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

FIGURE 247

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGGCGCAAGGGTGAGGGCGGCCCCAGAA
 CCCAGGTAGGTAGAGCAAGAAG**ATG**GTGTTTCTGCCCTCAAATGGTCCCTTGCAACCATG
 TCATTTCTACTTTTCCCTCACTGTTGGCTCTCTTAACCTGTGTCCACTCCTTCATGGTGTGAGAG
 CACTGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCTTGGAATAAAATACGACTTC
 CTGAGTACGTCATCCCAGTTCATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACC
 TTCTGGGGAACCACGAAAAGTAGAAATCACAGCCAGTCAGCCCACCAGCACCATCATCCTGCA
 TAGTCACCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGCTATCGG
 AAGAACCCCTGCAGGTCTTGGAACACCCCCCTCAGGAGCAAATTGCACTGCTGGCTCCCGAG
 CCCCTCCTTGTCGGGCTCCCGTACACAGTTGTCAATTCACATATGCTGGCAATCTTTGGAGAC
 TTTCCACGGATTTTACAAAAGCACCTACAGAACCAAGGAAGGGGAAGTGAAGGATACTAGCAT
 CAACACAATTTGAACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGATGAACCTGCCTTC
 AAAGCAAGTTTCTCAATCAAATTAGAAGAGAGCCAAGGCACCTAGCCATCTCCAATATGCC
 ATTGGTGAAATCTGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGATGTCACTGTGA
 AGATGAGCACCTATCTGGTGGCCTTCATCATTTTCAGATTTTGAGTCTGTGAGCAAGATAACC
 AAGAGTGGAGTCAAGGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGC
 ACTGGATGCTGCGGTGACTCTTCTAGAATTTTATGAGGATTATTTTCAGCATACCGTATCCCC
 TACCCAAACAAGATCTTGCTGCTATTTCCCGACTTTTCAGTCTGGTGCTATGGAAAACCTGGGGA
 CTGACAACATATAGAGAATCTGCTCTGTTGTTGATGCAGAAAAGTCTTCTGCATCAAGTAA
 GCTTGGCATCACAGTGACTGTGGCCCATGAACCTGGCCACCAGTGGTTTGGGAACCTGGTCA
 CTATGGAATGGTGGAATGATCTTTGGCTAAATGAAGGATTGCCAAATTTATGGAGTTTGTG
 TCTGTGAGTGTGACCCATCCTGAACTGAAAGTTGGAGATTATTTCTTTGGCAAATGTTTTGA
 CGCAATGGAGGTAGATGCTTTAAATTCCTCACACCCTGTGTCTACACCTGTGGAAAATCCTG
 CTCAGATCCGGGAGATGTTTGATGATGTTTCTTATGATAAGGGAGCTTGATTCTGAATATG
 CTAAGGGAGTATCTTAGCGCTGACGCATTTAAAGTGGTATTGTACAGTATCTCCAGAAGCA
 TAGCTATAAAAATACAAAAACGAGGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAG
 ATGGTGTAAGGGGATGGATGGCTTTTGCTCTAGAAGTCAACATTCATCTTCATCCTCACAT
 TGGCATCAGGAAGGGGTGGATGTGAAAACCATGATGAACACTTGGACACTGCAGAGGGGTTT
 TCCCCTAATAACCATCACAGTGAGGGGGAGGAATGTACACATGAAGCAAGCACTACATGA
 AGGGCTCTGACGGCGCCCCGGACACTGGGTACCTGTGGCATGTTCCATTGACATTATCACC
 AGCAAATCCAACATGGTCCATCGATTTTGTCTAAAAACAAAAACAGATGTGCTCATCCTCCC
 AGAAGAGGTGGAATGGATCAAATTTAATGTGGGCATGAATGGCTATTACATTGTGCATTACG
 AGGATGATGGATGGGACTCTTTGACTGGCCTTTTAAAAGGAACACACACAGCAGTCAGCAGT
 AATGATCGGGCAAGTCTCATTAAACAATGCATTTTCAGCTCGTCAGCATTGGGAAGTGTCCAT
 TGAAAAGGCCTTGGATTTATCCCTGTACTTGAACATGAAACTGAAATTATGCCCGTGTTC
 AAGGTTTGAATGAGCTGATTCCTATGTATAAGTTAATGGAGAAAAGAGATATGAATGAAGTG
 GAAACTCAATTTCAAGGCCTTCCTCATCAGGCTGCTAAGGGACCTCATTGATAAGCAGACATG
 GACAGACGAGGGCTCAGTCTCAGAGCAAATGCTGCGGAGTGAACACTACTACTCCTCGCCTGTG
 TGCACAACTATCAGCCGTGCGTACAGAGGGCAGAAGGCTATTTTCAGAAAGTGGAAGGAATCC
 AATGGAAACTTGAGCCTGCCTGTGCGACGTGACCTTGGCAGTGTGCTGTGGGGGCCAGAG
 CACAGAAGGCTGGGATTTTCTTTATAGTAAATATCAGTTTTTCTTTGTCCAGTACTGAGAAAA
 GCCAAATTGAATTTGCCCTCTGCAGAACCCAAAATAAGGAAAAGCTTCAATGGCTACTAGAT
 GAAAGCTTTAAGGGAGATAAAATAAAAACTCAGGAGTTTCCACAAATTCCTTACACTCATTGG
 CAGGAACCCAGTAGGATACCCACTGGCCTGGCAATTTCTGAGGAAAACTGGAACAACTTG
 TACAAAAGTTTGAACCTGGCTCATCTTCCATAGCCACATGGTAATGGGTACAACAAATCAA
 TTCTCCACAAGAACACGGCTTGAAGAGGTAAAAGGATTCTTCAGCTCTTTGAAAGAAAATGG
 TTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATGGAAGAAAACATCGGTTGGATGG
 ATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCTTGAACGTATG**TAAAAA**
 TTCCTCCCTTGCCCGGTTCTGTTATCTCTAATCACCAACATTTTGTGAGTGTATTTTCAA
 ACTAGAGATGGCTGTTTTGGCTCCAACCTGGAGATACTTTTTTCCCTTCAACTCATTTTTTGA
 CTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTTTCATGAATGGGCTA
 TCGCTACCATGTGTTTTGTTTCATCACAGGTGTTGCCCTGCAACGTAAACCCAAGTGTGGGT
 TCCCTGCCACAGAAGAATAAAGTACCTTATTCTTCTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 248

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH
 YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLOVLE
 HPPQEQIALLAPEPLLVLGPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA
 ARMAFPCFDEPAFKASFSEIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA
 FIISDFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAA
 IPDFQSGAMENWGTLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL
 WLNFGFAKFMFVSVSVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD
 DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDG
 FCSRSQHSSSSSSHHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD
 TGYLWHVPLTFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL
 TGLLKGTHTAVSSNDRASLINNAFQLVLSIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP
 MYKLMKCRDMNEVETQFKAFILRLRLDLIDKQTTWDEGSVSEQMLRSELLLLACVHNYQPCV
 QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSSTEKSQIEFALC
 RTQNKEKLQWLLDESFKGDKIKTQEFQILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGS
 SSIAHVMVGGTTNQFSTRTRLEEVKGFFSSLKENGSQLRCVQQTIETIEENIGWMDKNFDKIR
 VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

FIGURE 249

CAGCCACAGACGGGTCATGAGCGCGGTATTACTGCTGGGCCCTCCTGGGGTTCATCCTCCAC
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC
GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG
AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCGGCCCTCTCCCTGATC
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCCTG
TTGGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCAGTCTGCTTGTCTATGG
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCGTGGGTATGACTGAGA
ACTGCAATAGGAAAGATTTTCTGACCTGTCTATCGGGGGACCACCATTATGACACACGGAAAC
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT
GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG
GCTGCAGCACTGTTGGGGCTCAAATTTCCAGAAGACCACCATCCACTCAGCCCCTCCTGGG
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG
CAGCGTTCTGCTGAACTCCCTCCCTCCTCAAGCTGCCCCGTCCCAGGAGACCGGCAGTGTC
CTACCTGTGTGCAGCCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG
GGCGCCACTCATTTGTTATGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAT
GAGCATTCAGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCG
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGT
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG
GTGGGGAGTGTTTGGCCTTCCTGCTAACTCTATTACCCCCACGATTCTTCACCGCTGCTGA
CCACCCACACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC
CCATTCTGTCCATGAATCATCTTCCCCACACACAATCATTATATCTACTCACCTAACAGCA
AACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 250

MSAVLLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRRLRGGGIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSENTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVC
PSC

FIGURE 251

GCGACGGGCAGGACGCCCCGTTTCGCCTAGCGCGTGCTCAGGAGTTGGTGTCTGCTGCCTGCGCT
 CAGG**ATG**AGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG
 CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG
 CCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCCGGACGGCCTGGAAGAGTCG
 GCCCCACGGGAGAAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTGCTCAT
 GGAAAAATTGGTCCCATTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC
 TGGTCCTAATGGAGAACCAGGCCTCCCATGTGAGTGCAGCCAGCTGCGCAAGGCCATCGGGG
 AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTCGCC
 GGTGTGCGGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA
 CGCCCAGCTGTCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA
 ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCCTGTCTTCATCGGCATCAAC
 GACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA
 GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT
 CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG
 GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCATTGGGGGCCCCACATGTCCCTGCAGGGTT
 GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG
 GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG
 AAAGTGTTCCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA
 ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC
 TATACAATAAAAATCTTTAAGTAGTGCAGTAGTTAAGTCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 252

M R G N L A L V G V L I S L A F L S L L P S G H P Q P A G D D A C S V Q I L V P G L K G D A G E K G D K G A P G R P G R V G
P T G E K G D M G D K G Q K G S V G R H G K I G P I G S K G E K G D S G D I G P P G P N G E P G L P C E C S Q L R K A I G E
M D N Q V S Q L T S E L K F I K N A V A G V R E T E S K I Y L L V K E E K R Y A D A Q L S C Q G R G G T L S M P K D E A A N
G L M A A Y L A Q A G L A R V F I G I N D L E K E G A F V Y S D H S P M R T F N K W R S G E P N N A Y D E E D C V E M V A S
G G W N D V A C H T T M Y F M C E F D K E N M

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG
CACCAGTGTGTGAGGGGAGCAGGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACCA**ATG**AGGATCATGCTGCTATTACAGCCAT
CCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGG
TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCCTCGGCCC
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC
TTTAT**TAA**GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT
ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC
TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT
AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA
TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT
CCTACATTAAAAATATAATGTCTCTCTCTATTCTCAACAATAAAGGATTTTGCATATGAA
AA

FIGURE 254

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK
ALSQASTDPKESTSPEKRDMDHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK
SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGTGTGCCCGCTGTGCCTGTGTGCC
CGCGCTGTTCGCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG
GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC
CCATAGTTGCTGCAGGAGTGGAGCC**ATG**AGCTGCGTCCTGGGTGGTGTGCATCCCCTTGGGGC
TGCTGTTCTGCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCTCTTAGAGGAG
CTGCTCAGCAAATACCAGCACAAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA
GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT
CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCCGGGGGTGGGGC
CTGGGGCCACCAGCCTGCTCTGTTCCCCAGCCAGCTCTGTTCCCCAGCCAGTGCGTGTGATGG
CTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCGGCTCTGTTCTGTTTTGTTTGTGTTGTT
TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG
AAACCT**TAG**ACTCCCGGGGTAAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACCTACAG
GCATGCACCATGGTGCCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT
TGCCCAGGCTGGTCTTGAACCTCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG
CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATCTTGCCAAA
ACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC
TCCACTGGGAACACAGCTCTCAGCCTTTCCACCTGGAGGCAGAGTGGGGAGGGGGCCCAGGG
CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC
CTTAGCCCGTGTGAGCCTCACTTTCCACTTGAGAGTCTTTCCTCGCGTGGTTGCCATGACT
GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCTAGG
CTTTGCTAACC GGGAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT
GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG
TGGAAACTTCCTTCCTGGCTGGTTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG
GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT
TTTAAACAAACCACGTCCCAGCCTGGGTAAACATGGTAAAGCCCCGTCTCTACAAAAAAATC
CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG
GTGGAGGTGGGGGGTGGGAGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC
AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

FIGURE 256

MSCVLGGV IPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNEHSRVRRAIPREDKEEILML
HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR
GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

Downloaded from www.tandem.co.uk

FIGURE 257

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGG
GTCTGGGCTGCCCCTTGTCTCCTCCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGG
GTATGACTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGC
TTCCTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCAC
CCTCCACCATGCAAGATCTCAACACCATGTTGTCTGCAACACATTGACCAGCCATTGAAGCCTG
TGTCTTCTTGCCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTT
CAGCAGGCCCCCACCTCCTGAGTGGCAATAAATAAAATTCGGTATGCTG

258/330

FIGURE 258

MSGGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSTG
VTLHHARSQHHVCNT

7085550

FIGURE 259

AATTGTATCTGTGTAATGTTAAAACAAACGAAATAAAATAGAAGGAAAACTTTCTGAGTTT
CAAAAACAACAGACTAGTACTCTAAAGAACTCTTTAAAACAATTAAGTGTAGGATTGCAGT
TATGATTGGATATTATTTAATTCTGTTTCTGATGTGGGGTTCCCTCCACTGTGTTCTGTGTGC
TATTAATATTTACCATTGCAGAAGCTTCATTCAGTGTTGAAAATGAATGCTTAGTGGATCTG
TGCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAGTTCCCCTCCC
CTCCGATTGTTCTAAAT**TAA**TTGAAAGATGTCTGCTGTGGAAAAAGGCATGTATTTAAATCTG
TATGATTCTCAACCATCTTTAGTTGGGAAAGGTCCTTGAAAGCCAATGGAAATACTTTTTTT
TTTTCTTGGCACTAATCAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTA
GTAAATAGAAACCTGTGTTTATTCTCAGGTATTTTAGAAACAACAGCCATCATTTTATTTT
ATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGGGCTATCAAATATTACTTCAT
TCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTCTAGTTGCATTTTCTC
AGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTA
TCTTCAGATTACTTGATTCAAATAAACCAATTATGTTTGTAAATTGATATTAATAAAACCAGA
ATAAAAGTTCATATCTACCC

260/330

FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFVENECLVDLCLLRICYKLSGVPNQCRVPLP
SDCSK

Important features:

Signal peptide:

amino acids 1-29

Downloaded from www.jstor.org

FIGURE 261

GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGAGCACCACCGGAGCCCTTGAGACATCCTT
 GAGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTTTTGCAGG**ATG**ATGGTGGCCCTT
 CGAGGAGCTTCTGCATTGCTGGTTCCTTGCAGCTTTTCTGCCCCCGCCGAGTGTAC
 CCAGGACCCAGCCATGGTGCATTACATCTACCAGCGCTTTCGAGTCTTGAGCAAGGGCTGG
 AAAAATGTACCCAAGCAACGAGGGCATACTTCAAGAATTCCAAGAGTTCTCAAAAAATATA
 TCTGTCATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACCTT
 GGCCTGAGAGTTGAACGTGCCCAACGGGAGATTGACTACATAACAATACCTTCGAGAGGCTG
 ACGAGTGCATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAA
 GAGAAAAAGATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTC
 TTTGAAAATAGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATA
 ACTCTCCAAAGGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAAC
 ATACGGGCATTTCATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTC
 CTGGCAGGGAAACAGGCCAAGTGATCTACAAAGGTTTTCTATTTTTTTCATAACCAAGCAACTT
 CTAATGAGATAATCAAATATAACCTGCAGAAGAGGACTGTGGAAGATCGAATGCTGCTCCCA
 GGAGGGGTAGGCCGAGCATTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGT
 GGATGAGCATGGGCTCTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCA
 CAAAGATTGAGCCGGGCACACTGGGAGTGGAGCATTCATGGGATACCCCATGCAGAAGCCAG
 GATGCTGAAGCCTCATTCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCA
 GGGCCCTCATCGCATCACCTGCATCTATGATCCACTGGGCCTATCAGTGAGGAGGACTTGC
 CCAACTTGTCTTCCCCAAGAGACCAAGAAGTCACTCCATGATCCATTACAACCCAGAGAT
 AAGCAGCTCTATGCCTGGAATGAAGGAAACCAGATCATTTACAACTCCAGACAAAGAGAAA
 GCTGCCTCTGAAG**TAA**TGCATTACAGCTGTGAGAAAGAGCACTGTGGCTTTGGCAGCTGTTT
 TACAGGACAGTGAGGCTATAGCCCCTTCACAATATAGTATCCCTCTAATCACACACAGGAAG
 AGTGTGTAGAAGTGGAATACGTATGCCTCCTTTCCCAAATGTCACTGCCTTAGGTATCTTC
 CAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTTCAACAATGTCCATTACTCCCCAAA
 CTTCTGGCTCTCAAGGATGACCACATTCTGATACAGCCTACTTCAAGCCTTTTGTCTTACT
 GCTCCCCAGCATTTACTGTAACTCTGCCATCTTCCCTCCCACAATTAGAGTTGTATGCCAGC
 CCCTAATATTCACCACTGGCTTTTCTCTCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTT
 CAAATGTCTATTGATATTCTCCCATTTTCACTGCCCACTAAAATACTATTAATATTTCTTT
 CTTTTCTTTTCTTTTTTTTGTAGACAAGGTCTCACTATGTTGCCCAGGCTGGTCTCAAACCTC
 AGAGCTCAAGAGATCCTCCTGCCTCAGCCTCCTAAGTACCTGGGATTACAGGCATGTGCCAC
 CACACCTGGCTTAAATACTATTTCTTATTGAGGTTTAACCTCTATTTCCCCTAGCCCTGTC
 CTTCCACTAAGCTTGGTAGATGTAATAATAAAGTGAAAATATTAACATTTGAATATCGCTTT
 CCAGGTGTGGAGTGTTCACATCATTGAATTCTCGTTTACCTTTGTGAAACATGCACAAG
 TCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGTAACACAATTACAAAGTGAAAGATACAGC
 TAGAAAATACTACAAATCCCATAGTTTTTCCATTGCCCAAGGAAGCATCAAATACGTATGTT
 TGTTTACCTACTCTTATAGTCAATGCGTTTCATCGTTTTCAGCCTAAAAATAATAGTCTGTCCC
 TTTAGCCAGTTTTTTCATGTCTGCACAAGACCTTTCAATAGGCCTTTCAAATGATAATTCCTCC
 AGAAAACCAGTCTAAGGGTGAGGACCCCAACTCTAGCCTCCTCTTGTCTTGCTGTCTCTGT
 TTCTCTCTTTCTGCTTTAAATTCAATAAAAGTGACACTGAGCAAAAAAAAAAAAAA

FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL
QEAEKKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNPKVYLLIGSRNNTV
WEFANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKEPGLGVEHSWDT
PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH
YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQIRIRKEKPQQHNFTHRLLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPPK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTAAVSPCGRFVASC
FTPDKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AQETLKSIGALKK

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG
CAGTGTTTTGCCTTCACCCCAAGTGACCATGAGAGGTGCCACGCGAGTCTCAATCATGCTCC
TCCTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT
GGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT
GGGGCGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA
AGCACCACACCTGTCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTAC
CGCTGCTCCATGGACTTGAAGAACATCAATTTTTTAGGCGCTTGCCTGGTCTCAGGATACCCA
CCATCCTTTTCTGAGCACAGCCTGGATTTTTTATTTCTGCCATGAAACCCAGCTCCCATGAC
TCTCCCAGTCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAG
GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG
CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTA
AATGGCAGAAAGGACATTCCCCCTCCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG
CCCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCTTGGGCACAGGCTCTTGGGT
GCATTGCTCAGAGTCCCAGGTCTTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG
ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC
TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA
GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA
CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT
TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA
TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTCCTTAAACAACCTCTTTCCA
AGGATCAGCCCTGAGAGCAGGTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG
TGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGCACTGATTACAGACCAGGGAGG
CAACTACACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAA

FIGURE 266

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAI SLWLRGLRMCTPLGREGEECHP
GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

FIGURE 267

AGCGCCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAGGGGAGGCACCTTGAGAAATGTCTTTCTCCAGGACCCAAGTTTCTTCCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGCTGCTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGGCCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA

GCAAAGGAGCTATGGGAAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTCCTCTGTGCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCGTCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCAGCCTTATTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAAGTTCTATGGTCCACAAAGGCGGAAGATGATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACCTTCTTCCGAGCCTGGAACGGAGGCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTTCGTGGTGGGATCAGGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACTTTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAATGATGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTTATGGGATGTATTGTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTATTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGACAAAAATCTGAAAACTAATGAGGATTATTAAGCTAAACCTGGGAAATAGGAGGCTTAAATTTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCAGCACTTTGGGAGGCCAAGGTGAGCAAGTCACTTGAGGTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAAATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCACCACTGTATTCCAGCCTGGGTGACTGAGACTCTAACTAA

FIGURE 268

MSFLQDPSFFTMGMWSIGAGALGAAALALLLANTDVFLSKPQKALEYLEDIDLKTLEKEPR
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGWYNFFRAWNGGFSGNLEGEFILGGVFV
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

FIGURE 268

FIGURE 269

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCG
GGCCAGGTGCCCCGTGCGAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGA
AGCCCCCTTCCTCGGCGCTGCCAACCCGCCACCCAGCCCC**ATG**GC GAACCCCGGGCTGGGGCTG
CTTCTGGCGCTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACA
GACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATG
GCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTG
CTCCTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGG
CACCTACCGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGG
ACTCCAAGGAGACGGTGCAGGGCTGCCTGCCCATC**TAG**GTCCCCTCTCCTGCATCTGTCTCC
CTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAG
TGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGAGGATGGG
GCTATTCACTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

270/330

FIGURE 270

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIV
VFSLAALLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

10311103660

FIGURE 271

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTT
TTGGGATTTTAATTTTCAAACACAGCAGAATGACATTTTTTCTGTCACTATTATTATTGTTG
GTATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCT
ATCAAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTC
CTACCAAAGCTGTCAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGA
GGGTTAATTCTTGCTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAAACACTTAGATTCAATG
ATTGTAAATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATA
TATACAATAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACT
TTATTAATTTTTTAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGA
TCATATAATTTGATACAAATAAAAGAAAAGTGTTCTCTCCCCTTACAGAATTGACATTTTAA
ATGCGATACAGTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAA
AGAAGGGAAAATGTTGCCAAGGAAAAAAAAA

FIGURE 272

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK
GIVKGBNLDNRGLILGAEAWGRGVKKNT

FIGURE 273

GCCAGGAATAACTAGAGAGGAACA**ATG**GGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTTCT
 GTGCCTGCTGCACCAAGTCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA
 TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAATAATTGAACAAATAGAGGAT
 ATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGATTTTTTTTCAAAAA
 TGTATCTATATTAATTCCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG
 AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA
 TACACCAAGCAGTTCACAGAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCT
 TCTACTTGAAAAAAACAAAATGAATATGGACCACAGGCAAACCTGTTTGTCCATGAGTGGG
 CTCACCTCCGGTGGGGAGTGTTTGTATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG
 TCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA
 GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACCTGTATG
 GAAAAGATTGTCAATTCTTTCTTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG
 CAAAGTATTGATTCTGTTGTTGAATTTTGAACGAAAAAACCCATAATCAAGAAGCTCCAAG
 CCTACAAAACATAAAGTGCAATTTTGAAGTACATGGGAGGTGATTAGCAATTCTGAGGATT
 TTA AAAACACCATAACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATC
 AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT
 AAATCGAATGAATCAAGCAGCAAAACATTTCTGCTGCAGACTGTTGAAAAATGGATCCTGGG
 TGGGATGGTTCACTTTGATAGTACTGCCACTATTGTAATAAGCTAATCCAAATAAAAAAGC
 AGTGATGAAAGAAACACACTCATGGCAGGATTACTACATATCCTCTGGGAGGAACTTCCAT
 CTGCTCTGGAATTAAATATGCATTTTCAAGGTGATTGGAGAGCTACATTCCCAACTCGATGGAT
 CCGAAGTACTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTG
 AAACAAAGTGGGGCCATTGTTCATTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT
 AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTTCAGATGAAGCTCAGAACAATG
 GCCTCATTTGATGCTTTTGGGGCTCTTACATCAGGAATACTGATCTCTCCAGAGTCCGTT
 CAGCTCGAAAGTAAGGGATTAACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT
 TGATAGTACAGTGGGAAAGGACACGTTCTTCTCATCACATGGAACAGTCTGCCTCCAGTA
 TTTCTCTCTGGGATCCAGTGGAACAATAATGGAAAATTTACAGTGGATGCAACTTCCAAA
 ATGGCCTATCTCAGTATTCAGGAAGTGCAAAGGTGGGCACTTGGGCATACAATCTTCAAGC
 CAAAGCGAACCAGAAACATTAACATTTACAGTAACTTCTCGAGCAGCAAAATCTTCTGTGC
 CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATT
 GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCAT
 TGAATCACAGAATGGACATACAGAAGTTTTTGGAACTTTTGGATAATGGTGCAGGCGCTGATT
 CTTTCAAGAAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAATGGCAGATAT
 AGCTTAAAGTTTCGGGCTCATGGAGGCAAAACACTGCCAGGCTAAAATTACGGCCTCCACT
 GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAACCCGCCAA
 GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTTCAGCCGAACAGCATCCGGA
 GGTGCATTTGTGGTATCACAAGTCCCAAGCCTTCCCTTGCCCTGACCAATACCCACCAAGTCA
 AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG
 GAGATAATTTTGTGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT
 GATCTAAGAGACAGTTTTTGTGATGCTCTTCAAGTAAATACTACTGATCTGTCAACCAAGGA
 GGCCAACCTCCAAGGAAAGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCC
 ACATATTTATTGCCATTAAAAGTATAGATAAAAGCAATTTGACATCAAAGTATCCAACATT
 GCACAAGTAACCTTGTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC
 TACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTAT
 TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATT**TGA**ACCTTA
 ACGAAGAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTTAAAAACAAAACAATGTAAGT
 AAAGGATATTTCTGAATCTTAAATTCATCCCATGTGTGATCATAACTCATAAAAATAATT
 TTAAGATGTCGAAAAGGATACTTTGATTAAATAAAAAACACTCATGGATATGAAAACTGT
 CAAGATTAAAAATTTAATAGTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAAC
 AAAGATCCTTTTTTCACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTTCTGAAAT
 GATATTTCAAATTGCATCAAGAAATTTAAATCATCTATCTGAGTAGTCAAAATACAAGTAAA
 GGAGAGCAAATAAACACATTTGGA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 274

MGLFRGFVFLLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTY
 LFEATEKRFFFFKNVSILIPENWKENPQYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTEC
 GEKGEYIHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
 CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVE
 FCNEKTHNQEAPSLQNIKCNRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLV
 LDKSGSMGGKDRNLNRMNQAAKHFLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
 AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTGDGEDNTASSCIDEVKQSGAIVH
 FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
 LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG
 TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
 YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNMGVYSRYFTAYTENGRYSLKVRAGH
 GANTARLKLRPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRASGGAFVVSQV
 PSLPLPDQYPPSQITDLDATVHEDKIILTWTA PGDNFVDVGKVQRYIIRISASILDLRDSFDD
 ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLT SKVSNIAQVTLFIP
 QANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 275

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCG
GGCAGGGGTGACAACAGGTGTCTATCTTTTTGATCTCGTGTGTGGCTGCCTTCCTATTTCAAGGAAAG
ACGCCAAGGTAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACC
CCCAGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTT
GTGGTTGGAGGAGAGAACCCTTTGTGGGGCTGCGTTCTCTTAGCAGTGCTCAGAAAGTGAAGTTCCTGA
GGGTGGACCAGAAGAAAGGAAAGGTCCCCTCTTGCTGTTGGCTGCACATCAGGAAGGCTGTGATGGG
AATGAAGGTGAAAACCTTGGAGATTTCACTTCAGTCATGTGCTTCTGCCTGCAAGATCATCCTTTAAAA
GTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATG
CAAGCAGCTCCGGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAAGCCCTTCCGTGGGG
GCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTCGCC
GGGGCTGCTTGGCTGGATTTCCTGGGTGGTGGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
CCTGTACATGTTGGCCTGCACCCCAAAGGTGACGAGGAGCAGCTGGCACTGCCAGGGGCCAACAGC
CCCACGGGGAAGGAGGGGTACCAGGGCCGTCCTTCAGGAGTGGGAGGAGCAGCACCAGCAACTACGTGA
GCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGTGCAGGAGAGGAGTGAGCAGCTCAGGAA
TGGGCAGTACCAAGCCAGCAGTGTGCTGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
GCCGACCTCCTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGG
CCACAGAGTATGCAGCAGTGCCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGAGCTGG
CCTTACCCGCCACCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGGAAGCCATT
GAATCAGCCTTGGAGACCCTGAACCAATCCTGCAGAGAAGCAGCCCAATCACCCTTACACGGCCT
CTGATTTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCTTCAA
AGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTTCGACCATTTCAGCCCCATCATGAAAGTG
AAAAATGAAAAGCTCAACATGGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGG
ACAAGTTCCGGCAGTTTCATGCAGAATTTTCAGGGAGATGTGCATTTGAGCAGGATGGGAGAGTCCATCT
CACTGTTGTTTACTTTGGGAAAGAAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAA
GCTGCCAACTTCAGGAACCTTACCTTCATCCAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTG
ATGTTGGAGCCGCTTCTGGAAGGGAAGCAACGTCCTTCTCTTTTCTGTGATGTGGACATCTACTT
CACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAGGTATTTTATCCAGTT
CTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAAACAGC
AGCTGGTCATAAAGAAGGAAACTGGATTTTGGAGAGACTTTGGATTGGGATGACGTGTGATGATCG
GCAGACTTCATCAATATAGGTGGGTTTGTGATGTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCAC
CTTTATCGCAAGTATCTCCACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACC
TCTGGCATGAGAAGCGCTGCATGGACGAGCTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAA
GGCCATGAACGAGGCATCCACGGCCAGCTGGGCTGCTGGTGTTCAGGCACGAGATAGAGGCTCAC
ATTGCGAACACAGAAACAGAAAGCAAGTAGCAAAAAAACA**TGA**ACTCCAGAGAAGGATTTGTGGGAGA
CACTTTTTCTTTCTTTTGCATTTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAGGACG
ACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGTGTGGGCTTTTTTAC
AACAGAAATCAAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCTGACA
AAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTGTGATGGTGTGTACAATACACT
GAGACCTGTTGTTTGTGTGCTCATTGAAATATTCATGATTTAAGAGCAGTTTGTAAAAAATTTCAT
TAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGAGCCTATCAGCAGGGCTCTAGTTTCTAGG
AATGCTAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACTAGTGAATACATTAAAGTA
AAATAAAATGGACCAGAAAAGAAAAGAAACCATAAATATCGTGTGATATTTTCCCCAAGATTAACCA
AAAATAATCTGCTTATCTTTTTGGTGTGCTTTTAACTGTCTCCGTTTTTTTTCTTTTATTTAAAAAT
GCACTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTTGCAAGCCTTACAAGAGA
GCACAAGTTGGCCTACATTTTTTATATTTTTTAAAGAAGATACCTTTGAGATGCATTATGAGAACTTTCA
GTTCAAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAAT
GTCAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATACAGACGTACAGATACTTTCTCTGAA
GAGTATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTCTGCTTTACAGAA
AAGGAAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCACATTTTCTCCTCA
GAAGTAGGGACCGCTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCT
TTTCAAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATAT
ATATATATATATTTGTAAAGATCAATCCATCTGCCAGAACTAGTGGGATGGAAGTTTTTGTCTACAT
GTTATCCACCCAGGCCAGGTGGAAGTAACCTGAATTAATTTTTTAAATTAAGCAGTTCTACTCAATCA
CCAAGATGCTTCTGAAAATTGCATTTTATTACCATTTCAAACCTATTTTTTAAAAATAAATACAGTTA
ACATAGAGTGGTTTTCTTCATTTCATGTGAAAATTATAGCCAGCACCAGATGCATGAGCTAATTATCT
CTTTGAGTCTTCTGCTTCTGTTTGTCTCACAGTAACTCATTGTTTAAAGCTTCAAGAACATTCAAGC
TGTGGTGTGTTTAAAAAATGCATTTGATTGATTTGTACTGGTAGTTTATGAAATTTAATTTAAACAC
AGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

FIGURE 276

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPV RKDKRDELVEAIES
ALETNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSP
MKVKNEKLN MANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLN
TCRLNTQPGKKVFYPVLFSQYNPGIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDF
INIGGFDLDIKGWGGEDVHLYRKYLHSNLIVVRTPV RGLFHLWHEKRCMDELTPEQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

FIGURE 277

GAAAGA**ATG**TTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTATGCTGAACCTCTGTCAACC
AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT
ATGCCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA
GTTCCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT
ATCATTTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC
AATCAGCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAAACCT
CTGGAATTTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG
GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT
CAGGGATCTGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT
AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCCTCTGATCCCCCTGGACATGAAGGG
GGGCATATTAATGATGCCTTCA**TGA**CAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT
TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA
CCAAGAGCAGATCATATATTTTGTTCACCATTCTTCTTTTGTAAATAAATTTTGAATGTGCT
TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC
TCAAAATATTCTAAAATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG
TAGTTATTGATTTAAGCATTTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC
AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT
TGAAAATGGATCCTTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG
TGAGAAGTAATTATTGTAAATGGATGGATAAAAAATGGAATTACTCATATACAGGGTGAATT
TTATCCTGTTATCACACCAACAGTTGATTATATATTTTCTGAATATCAGCCCCTAATAGGAC
AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG
TAATAATCATCTCTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

278/330

FIGURE 278

MLWLLFFLVTAIHAE LCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVAFSMRKVP
NREATEISHVLLCNVTQRVSFWFVVTDP SKNHTLP AVEVQSAIRMNKNRINNAFFLNDQTLE
FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKC
ENMITIENGIPSDPLDMKGGILMMPS

0932041403560

FIGURE 279

AACTCAAACCTCCTCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG
 GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATGT**
 ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGG
 CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC
 TCGGTTAAAATGCACCTTCTCCAGCTTTGCCCCCTGTGGGTGATGCTCTAACAGTGACCTGGA
 ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC
 CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA
 TGCCTCCATCCTTCTCTGGAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA
 AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA
 CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT
 AATAGTAATTGTAGTGGTCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC
 ATAAAGTGGTGGAGATAAAATCAAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT
 GTTTATTTAGAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA
 GAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT
 TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC
 AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA
 GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAACAAGAAGCTACATTTTTTGCCCTTAA
 GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC
 AATTTGTCTGTTACATTTCTTTTACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA
 ATGTGTTTACTCTCTTTTCTTCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTG
 TTTCTGATTAAACAGTAAATCCTAAATTCAAACCTGTAAATGACATTTTTATTTTTATGTCTC
 TCCTTAACTATGAGACACATCTTGTTTTACTGAATTTCTTTCAATATTCAGGTGATAGATT
 TTTGTCG

280/330

FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT
WNRPLDGGPEQFVFYYHIDPFQPMSEGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ
VKNPPDVGVI GEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVL FQHYRKKRWAER
AHKVVEIKSKEERLNQEKKVSVYLEDTD

bioRxiv preprint doi: <https://doi.org/10.1101/140250>; this version posted April 11, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 281

GCATTTTTGTCTGTGCTCCCTGATCTTCAGGTCACCACCATGAAGTTCTTAGCAGTCCTGGT
ACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTCCAGCTG
ACACGTATCCAGCTACTGGTCCTGCTGATGATGAAGCCCCTGATGCTGAAACCACTGCTGCT
GCAACCACTGCGACCACTGCTGCTCCTACCACTGCAACCACCGCTGCTTCTACCACTGCTCG
TAAAGACATTCCAGTTTTTACCCAAATGGGTTGGGGATCTCCCGAATGGTAGAGTGTGTCCCT
GAGATGGAATCAGCTTGAGTCTTCTGCAATTGGTCACAACCTATTCATGCTTCCTGTGATTTC
ATCCAACCTACTTACCTTGCCTACGATATCCCCTTTATCTCTAATCAGTTTATTTTCTTTCAA
ATAAAAAATAACTATGAGCAACATAAAAAAAAAAAAAA

282/330

FIGURE 282

MKFLAVLVLLGVSI FLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTTAT
TAASTTARKDIPVLPKWVGDL PNGRVCP

protein database

FIGURE 283

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGACCCCATAGGGGA**ATGC**
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCCTGGTCCCTCCTTGCCGAGAGA
GTGTCCTGGGTCAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGGAC
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGG
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGCG
GCAAACCTCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGCTTTGTGCCAG**TAA**AACCACAGGCTGG
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTTTACCACATTAGCAATTAAACTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA

FIGURE 284

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHRLQPFVTERTLGKVQRWSGVHTQTGGRAG
GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

FIGURE 285

GTCA**ATG**CCAGTGCCTGCTCTGTGCTGCTCTGGGCCCTGGCAATGGTGACCCGGCCTGCCTCA
GCGGCCCCCATGGGCGGCCAGAACTGGCACAGCATGAGGAGCTGACCCTGCTCTTCCATGG
GACCCTGCAGCTGGGGCCAGGCCCTCAACGGTGTGTACAGGACCACGGAGGGACGGCTGACAA
AGGCCAGGAACAGCCTGGGTCTCTATGGCCGCACAATAGAACTCCTGGGGCAGGAGGTGAGC
CGGGGCCGGGATGCAGCCCAGGAACTTCGGGGCAAGCCTGTTGGAGACTCAGATGGAGGAGGA
TATTCTGCAGCTGCAGGCAGAGGCCACAGCTGAGGTGCTGGGGGAGGTGGCCCAGGCACAGA
AGGTGCTACGGGACAGCGTGCAGCGGCTAGAAGTCCAGCTGAGGAGCGCCTGGCTGGGCCCT
GCCTACCGAGAATTTGAGGTCTTAAAGGCTCACGCTGACAAGCAGAGCCACATCCTATGGGC
CCTCACAGGCCACGTGCAGCGGCAGAGGCGGGAGATGGTGGCACAGCAGCATCGGCTGCGAC
AGATCCAGGAGAGACTCCACACAGCGGCGCTCCCAGCC**TGA**ATCTGCCTGGATGGAACCTGAG
GACCAATCATGCTGCAAGGAACACTTCCACGCCCCGTGAGGCCCTGTGCAGGGAGGAGCTG
CCTGTTCACTGGGATCAGCCAGGGCGCCGGGCCCCACTTCTGAGCACAGAGCAGAGACAGAC
GCAGGCGGGGACAAAGGCAGAGGATGTAGCCCCATTGGGGAGGGGTGGAGGAAGGACATGTA
CCCTTTCATGCCTACACACCCTCATTAAGCAGAGTCGTGGCATTTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 286

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK
ARNSLGLYGRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK
VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ
IQERLHTAALPA

Downloaded from www.oxfordjournals.org/

FIGURE 287

GGCAAC**ATG**GCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAAATTTGCAT
 CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTAAAAGCCAGGA
 AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC
 TGGACAGAAGTCAATGCCTTGAAGGAAATTCAAGCCCTGCAGACAGTCTGTCTCCGAGGCAC
 TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCATGAGGCCAATG
 AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC
 CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA
 CATGGTCACGGAAGGCAAGTTTGTGACGTCAACGGAATCGCTATCTCCTTCCTCAACTGGG
 ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACGTGTGTCCTGTTCTCCCAATCAGCTCAG
 GGCAAGTGGAGTGATGAGGCCTGTCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC
 TAAAT**TAG**GTCTTTCTCCAATGTGTCCTCCAAGCAAGATTCATCATAACTTATAGGTTTCATGA
 TCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAATAAATTGCAACACAAGATCAAT
 GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT
 TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT
 AAACAGACTAAAATCTTTCTCTCTAGTCTTTCTCACTTGTACAAACCCAGTTTGTTTTCAA
 AAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAGCAAGCTTAGGCTACCTGAAAGATT
 TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCTACATCAGAGACTCTAGGT
 GCTATATAATCCAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC
 TTGTCAGCCCATACCCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT
 GCCATATCAGAACACAAACCCCTGAAGAGGTCTGATTTGATTTTTTTTTTTTTCTTCATGCC
 TACCCTTTTTTTTGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT
 CAATTTTCATTCCCACCATTCATTACAACCTCTAACTTAAATGGGTAAACCCTAAGGCATAT
 CAAAGAAGCAGATTGCATGATAAACGGAAATAGAAAAAAGAACCTACATTTATTTTGCTTT
 AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCCCTTTTTTACATTT
 TCGTATATTTATTTTTTTTTTAGCCATCATTATATGTTTAAAGTCTATTATGGGCAACCAATCTT
 TGGAAGCTGAAAACCTGAATTTAAAGAATGCTATCTTGGAATAATTGCATACGTCTGTGCAATT
 TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT
 TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTTAGAAGCAAACAATTTT
 AAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT
 CCCACTTTGCAAACCTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTTCATTGCTCA
 ATAATAAAGCCTGAATTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSKRRVRDKDGLKTQIEKLWT
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQ GK
WSDEACRSSKRYICEFTIPK

FIGURE 289

GCGAGGACCGGGTATAAGAAGCCTCGTGGCCTTGCCCGGGCAGCCGCAGGTTCCCCGCGCGC
CCCGAGCCCCCGCGCC**ATGA**AAGCTCGCCGCCCTCCTGGGGCTCTGCGTGGCCCTGTCCTGCA
GCTCCGCTGCTGCTTTCTTAGTGGGCTCGGCCAAGCCTGTGGCCCAGCCTGTCGCTGCGCTG
GAGTCGGCGGCGGAGGCCGGGGCCGGGACCCTGGCCAACCCCTCGGCACCCTCAACCCGCT
GAAGCTCCTGCTGAGCAGCCTGGGCATCCCCGTGAACCACCTCATAGAGGGCTCCCAGAAGT
GTGTGGCTGAGCTGGGTCCCCAGGCCGTGGGGGCCGTGAAGGCCCTGAAGGCCCTGCTGGGG
GCCCTGACAGTGTTTGGC**TGA**GCCGAGACTGGAGCATCTACACCTGAGGACAAGACGCTGCC
CACCCGCGAGGGCTGAAAACCCCGCCGCGGGGAGGACCGTCCATCCCCTTCCCCCGGCCCT
CTCAATAAACGTGGTTAAGAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAA

290/330

FIGURE 290

MKLAALLGLCVALSCSSAAFLVGS AKPVAQPVAALESAAEAGAGTLANPLGT LNPLKLLLS
SLGIPVNH LIEGSQKCV AELGPQAVGAVKALKALLGALT VFG

FIGURE 290

FIGURE 291

TGAAGGACTTTTCCAGGACCCAAGGCCACACACTGGAAGTCTTGCAGCTGAAGGGAGGCACT
 CCTTGGCCTCCGCAGCCGATCACATGAAGGTGGTGCCAAGTCTCCTGCTCTCCGTCTCCTG
 GCACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCTCAGTCGCCAGAGACCCAGCCCC
 TCAGAACCAGACCAGCAGGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGG
 CCAGCGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGGCAGCAGCTT
 GCCAAGGAGACTTCAAACCTTCGGATTGAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGG
 CAACATGGTCTTCTCTCCATTTGGCATGTCTTGGCCATGACAGGCTTGATGCTGGGGGCCA
 CAGGGCCGACTGAAACCCAGATCAAGAGAGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAG
 CCCGGGCTCCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAACCT
 GGGCCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCT
 TCAATTTATCCAAGAGGTATTTTGATACAGAGTGCCTGCCTATGAATTTTCGCAATGCCTCA
 CAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTTCCCAAACCT
 GTTTGATGAGATTAATCCTGAAACCAAATTAATTCTTGTGGATTACATCTTGTTCAAAGGGA
 AATGGTTGACCCCATTTGACCCTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTAC
 AAGACCATTAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAA
 TTTTCGTTGTCATGTCTCAAACCTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCTCA
 TGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACA
 TGGCTCAGAAACATGAAAACCAGAAACATGGAAGTTTTCTTCCGAAGTTCAAGCTAGATCA
 GAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCCTTTG
 CTGACCTTAGTGAACCTCTCAGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGA
 ACAGTGATTGAAGTTGATGAAAGGGGCACTGAGGCAGTGGCAGGAATCTTGTCAGAAATTAC
 TGCTTATTCCATGCCTCCTGTCATCAAAGTGGACCGGCCATTTCAATTCATGATCTATGAAG
 AAACCTCTGGAATGCTTCTGTTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTATAAATTCAGG
 ACATGCATAAGCACTTCGTGCTGTAGTAGATGCTGAATCTGAGGTATCAAACACACACAGGA
 TACCAGCAATGGATGGCAGGGGAGAGTGTTCTTTTTGTTCTTAACCTAGTTTAGGGTGTCTC
 AAATAAATACAGTAGTCCCCACTTATCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGC
 CTGAAACGGTGGACAGTGCTGAACCTTATATATATTTTTTCTTACACATACATACCTATGAT
 AAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAA
 TGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACCTGATTATAGAGAAGGCTA
 CTAAGTGACTCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAGAATTCA
 CATCCTGGGTGGGACAGAGCAGGACGATGCAAGATTCATCCCCTACTCAGAATGGCATGC
 TGCTTAAGACTTTTAGATTGTTTATTTCTGGAATTTTTCATTTAATGTTTTTGGACCATGGT
 TGACCATGGTTAACTGAGACTGCAGAAAGCAAAACCATGGATAAGGGAGGACTACTACAAAA
 GCATTAAATTGATACATATTTTTTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 292

MKVVPSLLLSVLLAQVWLVPG LAPSPQSPETPAPQNQTSRVVQAPREEEDEQEASEEKAGE
EEKAWLMASRQQ LAKETSNFGFSLLRKISMRHDGNMVFSPFGMSLAMTGLMLGATGPTETQI
KRGLHLQALKPTKPGLLP SLFKGLRETL SRNLELGLSQGSFAFIHKDFDVKETFFNLSKRYF
DTECVPMNFRNASQAKRLMNHYINKETR GKIPKLFDEINPETKLILVDYILFKGKWLT PFD P
VFTEVDTFHLDKYKTIKVPM MYGAGKFASTFDKNFRCHVLKLPYQGNATMLVVLMEKMGDHL
ALEDYLT TDLVETWLRNMKTRNMEVF FPKFKLDQKYEMHELLRQMGIRRI FSPFADLSELSA
TGRNLQVSRVLRRTVIEVDERGTEAVAGILSEITAYSMPPVIKVDRPFHFMIYEETSGMLLF
LGRVVNPTLL

FIGURE 293

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGAC**ATG**AG
GAGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAG
CACCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAG
GCCTGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTGTCTGTTCCC
TGTCCAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTTCAGGGCAGGGGGCCCCA
TCCTTCCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGTGTCCTGAGTCCC
GAGCCCGACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGGACCAGGGCGAGGAGAGGCC
CCGGTTGTGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACA
TCTACCACCCCCAG**TAG**GGCTCCAGGGGCCATCACTGCCCCCGCCCTGTCCCAAGGCCCAGG
CTGTTGGGACTGGGACCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCCAGCAGGCAAA
AAAAAAAAAAAAAAAAA

FIGURE 294

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL
FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE
RPRLWVMPNHQVLLGPEEDQDHIYHPQ

Downloaded from www.oxfordjournals.org/

FIGURE 295

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTG
TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATG**AACCAACTCAGCTTCCTGC
TGTTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAA
TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC
TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCT
GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG
CGTGGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCC
AGAGGGGGACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG
ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG
CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC
TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT
ATGGAGAAGGAAAGTGTGGACTGACAACGGCCCCGGTGATCCCTGTGGTCTATGATTTTGGC
GACGCCCAGAAAACAGCATCTTATTACTCACCTATGGCCAGCGGGAATTCAGTGCGGGATT
TGTTCAAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGG
TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT
CCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATATGGAACCATGTTGGTTA
CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TG**AGAGTTTTGTG
GGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCA
GTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAAA

FIGURE 296

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN
GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG
SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGLI
YQKYPVKYGEKGCWTDNGPVI PVVYDFGDAQKTASYSPYGQREFTAGFVQFRVFNNERAAN
ALCAGMRVTGCNTEHHCIGGGGYFPEAS PQQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLL
FYR

FIGURE 297

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC
 CACGAGGCTGCCGCATCCTGCCCTCGGAACA**ATG**GGACTCGGCGCGCGAGGTGCTTGGGCCG
 CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCGCCCATGAAAGCGCAGCC
 ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACCTCCAGTGCTAACTCAAC
 AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAAC
 CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG
 GCATCTAATAACAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC
 TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG
 TAACCCACAATAGTTTCAAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCAT
 TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC
 GCTGGGAGTTTTATCTATTTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC
 GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA
 ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAACAATATTCT
 CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA
 AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT
 TCATACAATGGTTTTAGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTGGCT
 GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA
 TGCCATCTGGGCATACAAATAAGAAGTTTGTACAGCACTCAGGATTTTGGGTATCTTTTGT
 AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA
 CAGAAATTATACAATCAAACCTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG
 TGCTTTAAACTGTAGTAGTTGGTCTAGAAACAAATACTCC

FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVP
SDH
TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSV
SQN
TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSI
LYIG
CKMYYSRRGIRYRTIDEHDAII

FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCTGAGCCGAGCCGGGAGCCGG
TCGCGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCAGCGGATGGCGACCCCTGTGGGGAGGC
CTTCTTCGGCTTGGCTCCTTGCTCAGCCTGTTCGTGCCTGGCGCTTTCCGTGCTGCTGCTGGC
GCAGCTGTCAGACGCCGCCAAGAATTTGAGGATGTCAGATGTAAATGTATCTGCCCTCCCT
ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT
CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA
ATGCAAATATGAAGAAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA
TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG
CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT
TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG
AATATGCACAGCAGCGCTGGAAGCTTCAAGTCCAAGAGCAGCGAAAGTCTGTCTTTGACCGG
CATGTTGTCCTCAGCTAATTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA
CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTTCATTTTAATACCTTGTTGATTTACCAACT
GTTGCTGGAAGATTCAAACTGGAAGCAAAACTTGCTTGATTTTTTTTTTTCTTGTTAACGTA
ATAATAGAGACATTTTTTAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG
TGACTTTTACTAATAAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA
AGCACTCTCTTTTTTACCACATAGTTTTTAACTTGACTTTCAAGATAATTTTTCAGGGTTTTTG
TTGTTGTTGTTTTTTTTGTTTGTTTGTTTTGGTGGGAGAGGGGAGGGATGCCTGGGAAGTGGTT
AACAACTTTTTTCAAGTCACTTTACTAAACAACTTTTTGTAAATAGACCTTACCTTCTATTT
TCGAGTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG
ACTTTTGCACCTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT
CTAAAATGCCTGGTGGCTTTTACAAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG
CAATGCATCCTAGAACAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG
GTGTGTGTGGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGACACT
TGCAATAAAGAAATTTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATATACATTTG
TCAGCATTTCCGGTTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT
AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT
TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAAATAATTTTTTGAACATCAA

5

5

FIGURE 301

GCACCTGCGACCACCGTGAGCAGTCATGGCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTT
CTGGGCTTGTCTCTGGCTCTGTCTGCTGCTGCTGCCCAAGGCCTTCCTGTCCCGCGGGAAGCGG
CAGGAGCCGCGCCGACACCTGAAGGAAAATTGGGCCGATTTCCACCTATGATGCATCATCA
CCAGGCACCCTCAGATGGCCAGACTCCTGGGGCTCGTTTCCAGAGGTCTCACCTTGCCGAGG
CATTTGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGGAAGAGGTCTG
ATGGGGCAGATTATTCCAATCTACGGTTTTGGGATTTTTTTATATATACTGTACATTCTATT
TAAGGTAAGTAGAATCATCCTAATCATATTACATCAATTGAAAATCTAATATGGCGATAAAAA
TCATTGTCTACATTAAACTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTA
AATCCTGCCTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTTCAATTTTACATAAGAATG
TTTACTCAATGTTTAAAGTGTTTTGCCCCAAAATTCACAACTAACAAGGCAGAACTAGGACTT
GAACATGGATCTTTTGGTTCTTAATCCAGTGAGTGATACAATTCAATGCACTCCCCTGCCA

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHQAPSDGQT
PGARFQRSHLAEAFKAKGSGGGAGGGGSGRGLMGQIIPYGFIFLYILYILFKVSRIILI
ILHQ

030304.430
T03T" T03650

FIGURE 303

CGGCTCGAGTGCAGCTGTGGGGAGATTTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT
GGATTTGAAAGTTGAGAGCAGCATGTTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTAC
TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG
GGTGATTACAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT
AGACTGGACTCTGTCAACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA
ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC
AATGATGGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA
AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG
AGGAGCCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTAGATGGGATGTGTTTTCCAG
AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA
GGAGATTGTATTTCTGTTACTACCACAAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG
GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG
CTTCAAGGAGTGAGGGAGTCAGATGGAGGAAACTACACCTGCAGTATCCACCTAGGGAACCT
GGTGTTCAAGAAAACCATTTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCC
CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC
TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA
GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG
AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA
CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA
CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACTCACTTGAAAAAAGTCAGGTGGGG
GAATGCCAAAAACACAGCAAGCCTTTTTGAGAAAGAATGGAGAGTCCCTTCATCTCAGCAGCGG
TGGAGACTCTCTCCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTTCAGACTCCCGCTCTC
CCAGCTGTCCTCCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG
CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC
CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTT
GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA
AAAACCAACCCAAATCAA

FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCI FKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGT YICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKLRMSVEYSQSWG HFQNRVNLVGDI FRNDGSIMLQGVRES DGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVT PAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKTCGNKSSVNSTV
LVKNTKKTNP EIKEKPCHFERCEGEKHIYSPIIVREVIEEEE PSEKSEATYMTMHPVWPSLR
SDRNNSLEKKSGGGM PKTQQAF

T05T T405660

FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG
GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG
AAACTACATTTTGGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CA
GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG
GCCCTGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG
ATGGTTGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA
TGAGAATGAAAATCGCACAGGAACCTTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG
TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCCTGTGACACAAAC
TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAGAGAG
TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG
AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT
GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGAAGA
TGGAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG
AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT
TAATGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT
ATGAATGCATCAGTAGCTGAAAAAAAAAAAAA

FIGURE 306

MQDEDGYITLNIKTRKPALVSVGPASSSWVRVMALILLILCVGMVVGLVALGIWSVMQRNYL
QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLWE
ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEEFL
EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

CSGTT 1408650

FIGURE 307

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGC
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG
CGGAGGAGAAGGAGGAGGAGGCGAACCAGAGAGGGGCAGCAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTGAGCAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTCACACCTGAGTGCA
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC
G**TAG**CCAGTGAGGGCAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGG
CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFSGSKRRRRRP
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA
MNSEGILYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMSHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

FIGURE 309

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG
ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG
CAACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCT
CCATTTAGAGTGTAGCAAAGGAAAAAACACCAAGGTTGGGTTCCCTTCCTGACATTGGCAGTG
CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTAC
AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAGGTGCCTGAAGATATTTAA
ACCACGTCTTGAAATTTAGTGGGTCTTGGCTTTGGGATAGGTGAAGTGAGGACAGACACTG
GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAACTCGAGGGTGGGATCCACTGAGG
AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCT
GTGGAGGGGGGTACGTGAGGGGGGGGTCTGGGGCTTATCCTCAGGTCCTGTGGGTGGGGCAG
CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG
CGCGCTCCGGGCGCCTGCCGGTTTGGGGGTGTCTCCTCCGGGGCGCT**ATG**GCGGCGCTGGC
CAGTAGCCTGATCCGGCAGAAGCGGGAGGTCCGCGAGCCCGGGGGCAGCCGGCCGGTGTTCGG
CGCAGCGGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG
CTGTCCAAGGTGCGACTGTGCGGGGGGCGGCCCGCGCGGCCGGACCGCGGCCCGGAGCCTCA
GCTCAAAGGCATCGTCACCAAACCTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCC
ACGGAAGCATCCAGGGCACCCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT
GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC
TGAGGGACTGCTCTACAGTTCGCCGCATTTACAGCTGAGTGTGCTTTAAGGAGTGTGTCT
TTGAGAATTACTACGTCTGTACGCCTCTGCTCTCTACCGCCAGCGTCGTTCTGGCCGGGCC
TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAA
GGCAGCTGCCCACTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC
ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCC**TGA**AATGTAGTCCCTGGACTG
GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACAACCTGT

FIGURE 310

MAALASSLIRQKREVPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR
GPEPQLKGIVTKLFCRQGFYLQANPDGSIQGTPEDTSSFTHFNLI PVGLRVVTIQSAKLGHY
MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR
VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

FIGURE 311

ATGGCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTG
GGACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACG
GCAACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGG
CGCCAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTT
GCAAATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCT
TCAACCTCATAACAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTAT
ATAGCCATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTACCCCTGAATGCAAGTT
TAAAGAATCTGTTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGG
AATCTGGTAGAGCCTGGTTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGA
GTAAAGAAAACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCG
AGAACCATCTTTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAA
GCACAAGTGCCTCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWD RPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLR
RQDPQLKGIVTRL YCRQGYLQMHPD GALDGTKDDSTNSTLFNLI PVGLRVVAIQGVKTGLY
IAMNGEGYLYPSELF TPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETV PKPGVTPSKSTSASAIMNGGKPVNKS KTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

FIGURE 313

GGGAGAGAGGAATTGACCATTGTAAAAGGAGACTTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAATG
GAAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGA
ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG
AAATAAACAGAGTTAGACCCGCGGGGGTTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCC
CTCCCCACCCCCAAAAAAAAGGATGATTGGAATGAAGAACCAGGATTACAAAGAAAAAAGTATGTTCATTTT
TTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAA
AGAACTGGTGTGGTGGTGTTTTTCTTTCTTTTTGAATTTCCCACAAGAGGAGAGGAAATTAATAATACATCTGC
AAAGAAATTTTACAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCA
CAGTTGGATTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTT
TTTTTAAATTTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCT
GGATGTTGCTGTGATCAGTCTGAAATACAACCTGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATG
TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCTAGGTTTAAACAGGGCCCTATTTGACCCCT
GCTTGTGGTGTCTGCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGT
GCTCCTGCAGCAACCAGTTTACAGCAAGGTGATTTGTGTTGCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCC
ACCAACACACGGCTGCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCAGCTTGAG
GCAGTTGGAAATCTTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGA
ACCTCAACACTCTGGAACCTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAA
CTGAAGGAGCTCTGGTTGCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAAACAGAATTCTTCTTT
GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAAC
TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAACTAGATGAG
CTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTTCAGGGTTTGATGCACCTTCAAAAAC
GTGGATGATACAGTCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCA
ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGCCATCATCTAGAGCGGATACAT
TTACATCACAACCTTGGAACTGTAACCTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTC
GAACACAGCTTGTTGTGCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA
ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT
GAGCTGAAATGTGGGCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAATGGAACAGTCATGACACA
TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTACAAATGTAACCTGTGCAAGATA
CAGGCATGTACACATGTATGGTGAGTAATTCCGTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCA
GCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACG
GACCACAGATAACAATGTGGGTCCCCTCCAGTGGTTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACAC
CACAGAGCACAAGGTGCAGACAGAAAACTTACCATCCAGTGAATGATATAAACAGTGGGATCCAGGAATT
GATGAGGTCATGAAGACTACCAAAATCATCATTGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT
GGTCATTTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAACCATCACGCCCCAACAAAGGACTGTTGAAATTA
TTAATGTGGATGATGAGATTACGGGAGACACCCCATGGAAGCCACCTGCCCATGCCTGCTATCGAGCATGAG
CACCTAAATCACTATAACTCATACAAATCTCCCTTCAACCACACAACAACAGTTAACACAATAAATTCATACAC
CAGTTCAGTGCATGAACCGTTATTGATCCGAATGAACTCTAAAGACAATGTACAAGAGACTCAAATCTAAACA
TTTACAGAGTTACAAAAAACAAACAATCAAAAAAAGACAGTTTATTAATAAATGACACAAATGACTGGGCTAA
ATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAACAAAAAGAAAAGAAATTTATTTATTAATAAATTCATATG
TGATCTAAAGCAGACAAAAA

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
 VRKNLREVPDGI STNTRLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
 NLNTLELFDNRLTTIPNGAFVYLSKLELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
 YISEGAFEGLSNRLRYLNLAMCNLREIPNLTPLIKLDELDSLGNHLSAIRPGSFQGLMHLQKL
 WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLHNPWNCNCDIL
 WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
 LKCRASLTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVNSVGN
 TTASATLNVTAATTTTFFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
 STRSTKFTFTIPVTDINSIGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHRQN
 HHAPTRTVEIINVDDIEITGDTPMESHLMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS
 VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419,
 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
 391-397, 422-428, 433-439, 531-537

FIGURE 315

GCGCCGGGAGCCCATCTGCCCCAGGGGACGCGGGCGCGGGGCGGGCTCCCGCCCGGCACAT
 GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTGGA
 GCGCCCGGGCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC
 GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCA
 CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACCTTGCCCTGCCACCATCAACTGGGGC
 TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
 GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
 AGTGGCCTTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGCAATTGAACCTCTGAAGC
 CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCAT
 GTCATCTTAAAAGTCTTAGTGAGACCATCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
 AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCATTGTGTATT
 ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
 GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
 CCAGTGCACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCAGTAAGTGTACAGT
 ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
 ATTTTCTCCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
 GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
 CCTCTTCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCCTCCACTCGCTCCACAGCAAAT
 AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
 CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
 CTAATCTGACCAAAGCAGAAACACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
 ACGGTCT**TGA**ATTACAATGGACTTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTC
 TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
 GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCTTATACAATACCAAACAAGCAAA
 AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
 AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
 AGGTGAATATACCTAAAACCTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT
 TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACCTTATTGGATT
 ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
 TGAGCTAACCACCTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
 TTCATTTGTGATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
 AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC
 TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
 ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT
 CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
 TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
 CCCAACATACCATTATAGTCTCTTCTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
 GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
 TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
 ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

316/330

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 317

CGCGAGGCGCGGGGAGCCTGGGACCAGGAGCGAGAGCCGCTACCTGCAGCCGCCGCCACGGCACGGCAGCCA
 CCATGCGCTCCTGCTGTGCTTCGTGCTCCTGTGCGGAGTAGTGGATTTCGCCAGAAGTTTGAGTATCACTACT
 CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA
 CCAGGGACCGCTGGACATCGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTAT
 ATTCTGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC
 AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACGTGTGAGATATTGGCACATATCAGTGCAAAGTGAA
 AAAAGCTCCTGGTGTGCAAATAAGAAGATTCATCTGGTAGTTCTTGTAAAGCCTTCAGGTGCGAGATGTTACG
 TTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAAATGTGAACCAAAAGAAGGTTCACTTCCATTACAG
 TATGAGTGGCAAAAATTTGTCTGACTCACAGAAAATGCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTAT
 ATCTGTAAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAACAGAGTGGGCTCTGATC
 AGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAAC
 TTGCTTGCTCTAGCGCTCATTGGTCTTATCATCTTTTGTGCTGTAAGCGCAGAGAAGAAAAATATGAAAA
 GGAAGTTCATCACGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG
 GCAGTAATCATTATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA
 GTACCAAGTGAAGACTTTGAACGCACTCCTCAGAGTCCGACTCTCCACCTGCTAAGTTCAGTACCCTTACAA
 GACTGATGGAATTACAGTTGTATTAATATGGACTACTGAAGAACTGAAGTATTGTATTATTTGACTTTATTTT
 AGGCCTCTAGTAAAGACTTAAATGTTTTTAAAAAAGCACAAGGCACAGAGATTAGAGCAGCTGTAAGAACAC
 ATCTACTTTTATGCAATGGCATTAGACATGTAAGTCAGATGTGATGTAAGTATAGTACGAGCCAAATTTTGT
 TAAAAAACCTATGTATAGTGACACTGATAGTTAAAGATGTTTTATTATATTTTCAATAACTACCACTAACAA
 ATTTTAACTTTTCATATGCATATTCTGATATGTGGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTTCAA
 AGGAAATTTTAAATTTCTTACGTTCTGTTTAAATGTTTTGCTATTTAGTTAAATACATTGAAGGGAATACCCG
 TTCTTTTCCCTTTTATGCACACAACAGAAACACGCGTTGTGATGCCTCAAACATTTTTTTATTTGCAACTACA
 TGATTTTACACAATTCTCTTAAACAACGACATAAAATAGATTTCTTGTATATAAATAAATACATACGCTCCA
 TAAAGTAAATTTCTAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTCTCGTATCCAACAGAGTTGATGC
 ACAATATATAAATACTCAAGTCCAATATTAATAAATTTAGGCCTTGACTAACTTTAATAAAATTTCTCAAAC
 TATCAATATCTAAAGTGATATATTTTTTAAAGAAAGATTATTTCTCAATAACTTCTATAAAAAATAAGTTT
 TTTGGCCCATCTAACTTCACTACTATTAGTAAAGAACTTTTAACTTTTAAATGTGTAGTAAGGTTTATTCTACCT
 TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAAATCCAGT
 TCCGGTCACAATGCATTCCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTTATTTTTTCTTCCC
 TGATTTGGATAACCAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCTCGATACATTCTGGCTT
 TTTTCTGGGCAAAGGGTGCCACATTGGAAGAGGTGGAAATATAAGTTCTGAAATCTGTAGGGAAGAGAACAT
 TAAGTTAATTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAGACAAAGTTACCCACAACACT
 GAGATCACATCTAAGTGACACTCCTATTGTGAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA
 TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCTAAACAA
 CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAAGGACATGCTTGTTTTAGTCCAGTGGTTT
 CCACAGCTGGCTAAGCCAGGAGTCACTTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT
 AGCAAACATATGCAGAAACAGAAAATCAACTACCGCATGTTCTCACTTATAAGTGGGAGGTAATGATAAGAACT
 TATGAACACAAAGAAGGAAACAATAGACATTGGAGTCTATTTGAGAGGGGAGGGTGGGAGAAGGAAAGGAGCA
 GAAAGATAACTATTGAGTACTGCCTTCACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA
 TGTTTACCTATGGAACAAACCTTCATGTGTATCCCTAAACCTAAATAAAAGTTAAAAA
 AA
 AAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK
APGVANKKIHLVVLVKPSGARCIVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSCTYTVRNRVGSQCLLRNVVPPSNKAGLIAGAIIGTLL
ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247,
245-251, 296-302

Myelin P0 protein.

amino acids 96-125

TGAAATGACTTCCACGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGA
ATGGTGAAGGTGCCTGTCTAACTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCC
CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTTCGC
CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG
AGAGATGAATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC
CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG
CTTCTCTGGAGCCAGGTATCAGGGGCCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT
GAAGGGGGTTGTTCCCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG
CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT
GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACGTTTTTCAAAAA
CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAAC
TTGTTCTCATCGTGTCAAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC
AGTGACACAGGCGGTTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC
TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC
TCTGAATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCTGTGTCAATTTCA
AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATT
TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG
GTGCCTCTGGATGCTGTGAAGAGTCTACAGAGAAGATTCTTGTATTTATTACAACCTCTATTT
AATTAATGTCAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG
CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAAATCCTTGCCACAGTGTGGGGCAG
TGGATGGGTGCTTAGTAAGTACTTAATAAACTGTGGTGCTTTTTTTTGGCCTGTCTTTGGATT
GTTAAAAAACAGAGAGGGATGCTTGGATGTAAAACTGAACTTCAGAGCATGAAAATCACACT
GTCTTCTGATATCTGCAGGGACAGAGCATTTGGGGTGGGGGTAAAGTGCATCTGTTTGAAAAG
TAAACGATAAAATGTGGATTAAAGTGCCCAGCACAAAGCAGATCCTCAATAAACATTTTCATT
TCCCACCCACACTCGCCAGCTCACCCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTTT
TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC
ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCTTGCTAATAAAAGACAACATAA
CTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

320/330

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFPCPLLATASQMOMVVLPCLGFTLLLWSQVSGAQQQEFHFGPCQVK
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLKTVFKNH
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL
TKALGEVDILLTWMQKFKYKL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

Downloaded from www.sagepub.com at [IP address] on [date]

FIGURE 321

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGC**ATGA**AAGTTACAGTGTGTTTCCCTTTGGCTC
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTC
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA
TCAGGAGCCAAACCCCAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
AATGCCACCAGAGTCATCCATGACAACCTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCT**TGA**TGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
GTGGGAGACAGCCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGACAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGCAAAAAGTCTACTGTGGTAT
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCCATATTTTACCTATGA

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ
RQCHCRQEATNATRVIHNDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG
AACGCGGCTACAATTAATACATAACCTTATGTATCATAACATACGATTTAGGTGACACTAT
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTC
GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACGATGTTGGGGGCCCCGCCT
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC
CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA
TGAGCAGAAGATACCTCTGCATGGATTTTCAGAGGCAACATTTTTGGATCACACTATTTTCGAC
CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC
TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC
CACCCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACC
CCCATACCACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT
GCTGAAGCCCCGGGCCCCGGATGACCCCGGCCCCGGCCTCCTGTTTCACAGGAGCTCCCGAGCG
CCGAGGACAACAGCCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTTCGAGTGAAC
ACGCACGCTGGGGGAACGGGCCCCGGAAGGCTGCCGCCCTTCGCCAAGTTCATCCTAGGGTCTG
CTGG

0998041.11501

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSSWGGLIHLYTATARN SYHLQIHKN GHVD
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGY
DVIHSPQYHFLVSLGRAKRAFLPGMNPPYPYSQFLSRNEIPLIHFNTPIPRRHTRSAEDDSE
RDPLNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNT HAGGTGPEGCRPFA
KFI

Important features of the protein:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

FIGURE 325

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG**
 GGCTGTCTCTGGGGTCTGGCTCTGCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG
 GAGCTCTGCAGGCCCCAGCACCCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG
 TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG
 ACCTCTTCTAGGGCCTCAACCCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA
 GAGAATTTCCCTTGCAAGAGAGACCAGGAGTTTCACAAAACATCTCCCAACTTCATGGTGC
 TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC
 ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA
 TGACAGCTCTGAAGAGGCAAAGACACTCACAAATGGACATATTGACATTGGCTCACACCTCCA
 CAGAAGCTAAGGGCCTGTCTCAGAGAGCAGTGCCTCTTCCGACGGCCCCCATCCAGTCATC
 ACCCCGTACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC
 GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT
 GGTCCCCGGGATCTGATGTCACCTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG
 GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA
 CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC
 CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC
 ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA
 CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA
 CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC
 AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTCAGCAGTGGGCAAAACAACCTTC
 CTTTGCTGGGAGCTCTGCTTCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC
 CTTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT
 CTTCCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA
 GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC
 GAGGCCGACCACAGACG**TGAG**TGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG
 TGGCTTCCCCGGAAGACCTCACTGACCCACAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC
 CGGGAACCTCCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA
 ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCTAGCCTG
 GGCCCCCACCAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG
 AAGGGCAGCATGTCCAAGCCCCTAACCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC
 CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCCAGAGGTGTCCTTGACTCACCTTGG
 CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA
 TTAAATTTCACTCAGTGTGGCCCAAAAAA

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTL
 ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
 TTVQTITGSDPEEAI FDTLCTDDSSEEAKTLTMDILT LAHTSTEAKGLSSESSASSDGPHPV
 ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTL LAEALVTVTNI
 EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
 STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLSGALVTVSRNPLEETSALSVETPSY
 VKVSGAAPVSI EAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPSTRD
 PLPSVPPTTTNSSRGTNSTLAKITTS AKTTMKPQQPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169,
 246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339,
 404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-
 320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 327

GCGGAGCATCCGCTGCGGTCCTCGCCGAGACCCCCGCGCGGATTGCGCCGGTCCTTCCCGCGG
 GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGCGCCGGGGTCTCTCGACGCCA
 GAGAGAAATCTCATCATCTGTGCAGCCTTCTTAAAGCAAACCTAAGACCAGAGGGAGGATTAT
 CCTTGACCTTTGAAGACCAAACTAACTGAAATTTAAAATGTTCTTCGGGGGAGAAGGGAG
 CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT
 GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTTCTAAGGGAATC
 AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC
 AAAAAACATATCAGGGGACAAAGCATGTAACCTTGATGATCTTCGACACTCGAAAAACAGCTA
 GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA
 AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG
 CCAAGAGTTACCCAGGAAGATTCTCTCTTACATGGCCAATTTTCAACAGCAGTCACTCCCC
 TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT
 CAGAAGTTTGGATCCTCAGATCACCTGGAGAACTATTTAAGATGGATGAAGCAAGTGCCCA
 GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTTACAATTTTCTCTGATCAAGAAA
 TAGCTCATCTGCTGCCTGAAAATGTGAGTGCCTCCAGCTACGGTGGCAGTTGCTTCTCCA
 CATACCACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC
 TTCTGGGACTTCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC
 AGCCTCCCACGACCCTCATTTCTACAGTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG
 GCTACAACAGCAGTTCTGACTACCACCTTTGAGGCACCTACGGACTCGAAAGGCAGCTTAGA
 AACCATAACGTTTACAGAAATCTCCAACCTTAACCTTTGAACACAGGGAATGTGTATAACCCTA
 CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAACTGCTTCCTGGGAAGGT
 AGGGAGGCCAGTCCAGGCAGTTCCTCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC
 ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCTGTTTCTGCTGATAGGCC
 TCGTCCTCCTGGGTAGAATCCTTTGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT
 TATTTGATCAATGGGATCTATGTGGACATCTAAGGATGGAACCTCGGTGTCTCTTAATTCATT
 TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG
 TATTTTGAAGACAGGAAAATGCCCCCTTCTGCTTTCCTTTTTTTTTTTGGAGACAGAGTCTT
 GCTCTGTTGCCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC
 CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA
 CCACACCTGGGTGATTTTTGTATTTTTTAGTAGAGACGGGGTTTACCAGTGTGGTCAGGCTG
 GTCTCAAACCTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG
 CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTGGTTTTTGGAGAAGGAATGAAGTG
 GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAATATTAGCTAAAAACAAAGCTCT
 ATGTAAAGTAATAAAGTATAATTGCCATATAAATTTCAAATTCAACTGGCTTTTATGCAAA
 GAAACAGGTTAGGACATCTAGGTTCCAATTCATTCACATTCTTGGTTCCAGATAAAATCAAC
 TGTTTATATCAATTTCTAATGGATTTGCTTTTCTTTTTATATGGATTCTTTTAAACTTATT
 CCAGATGTAGTTCCTTCCAATTAAATATTTGAATAAATCTTTTGTTACTCAA

FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410
><subunit 1 of 1, 431 aa, 1 stop
><MW: 46810, pI: 6.45, NX(S/T): 6
MFFGGEGSLTYTLVVICFLTLRLSASQNCLKKSLEDVVIDIQSSLKGIKRGNEPVYTSTQED
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASAQLLAYKEKGHSQSSQFSSDQEIALLPENVSALPATVAVASPHHTSATPKPATLL
PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP
TDSKGSLETIPFTEISNLTNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV
PENQYGLPFEEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSLDYLINGIYVDI

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

FIGURE 329

CTCCACGGTGTCCAGCGCCCAGAAATGCGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
 CCCAGGTTATGAAGCCCTGGAGGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
 CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
 GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
 GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
 ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
 TCTTTACTGATCTCTCTGTTTCGTCTTTCCAGGACCCTGCTGTCTCCCTCCCTTCTCCCAC
 CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC
 CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
 GCTGAGGCCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
 AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCCATGCAGC
 TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
 GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGGTGTGCTGAGCCTTCTGTGAGC
 CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
 CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
 CCTTCCCAGGCCCCCTGAGGGGGACGTGATCTCGATGCCTCCCTCCACACATCTGAGGAGGA
 GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
 GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAAGCTTCCACCTCAGCCTCAGAG
 TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGATGTTCCA
 GCCTGACCTAGAAGCGTTTGTACGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
 GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
 GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACCTCCTGGGC
 CTCATGCCCAGTGTCGGACCCTGCCTTCTCCCACTCCAGACCCACCTTGTCTTCCCTCCC
 TGGCGTCTCAGACTTAGTCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
 GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
 GTGAAAAACGTGATTCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG
 GACTCTGAATTCTAACAATGCCCAGTGAAGTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTG
 ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
 CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
 TCCAGGCCTTGGTCAGGTGAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
 TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCCTTTGGAAAAAATGATGAAGA
 AAACCTTGGCTCCTTCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTCTGGTGGCTAGAGA
 GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
 ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA
 GCACAATACTATTTTTTTTTCTTTTTCCATTATTATTGTTTTTAAAGACAGAATCTCGTGCT
 GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGGGTTCAGTGATT
 CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTGGCCAGGCTGGTCTTGAACCTCTGAC
 CTCAAATGAGCCTCCTGCTTCACTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
 TCTGGCCCTATTTCTTTTAAAAAGTGAATTAAGAGTTGTTTCAAGTATGCAAACTTGGAAG
 ATGGAGGAGAAAAAGAAAAGGAAGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
 TATTTTCGTTTTGTGTACTTCCCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTCTT
 TTTACAGAGCAATTATCTTGTATATACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
 ATCACTTTATTCAGCACTTCTCTGTGTTTTACAGACCTTTTATAAATAAAATGTTTCATCA
 GCTGCATAAAAAAAAAAAAAA

330/330

FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALLEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAE EEGQETMKGRVSI RDSRQELSLIVTLWNLTLDAGEYWCGVEKRGPD ELLISLFV
FPGPCPPSPSPSTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAE EKEAPSQAPEGD
VISMPPLHTSEEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

1036660